

Universitatea „Babeș-Bolyai”, Cluj-Napoca
Facultatea de Studii Europene
Școala Doctorală „Paradigma Europeana”

Anexă la Regulamentul de funcționare a Școlii Doctorale „Paradigma Europeană”

GHID DE REDACTARE A TEZEI DE DOCTORAT

2018

CUPRINS

CONCEPEREA, STRUCTURA ȘI PLANIFICAREA UNEI TEZE DE DOCTORAT	3
DEFINIREA, TEMATICA ȘI SPECIFICITATEA UNEI TEZE DE DOCTORAT	3
TITLUL TEZEI DE DOCTORAT	4
DIMENSIUNEA ȘI STRUCTURA TEZEI DE DOCTORAT	4
(I) DIMENSIUNEA TEZEI DE DOCTORAT	4
(II) STRUCTURA TEZEI DE DOCTORAT	5
ROLUL COORDONATORULUI	6
CONCEPEREA ȘI PLANIFICAREA TEZEI DE DOCTORAT	7
CONCEPEREA TEZEI DE DOCTORAT	7
ETAPE ÎN REDACTAREA TEZEI DE DOCTORAT	8
REGULI, RECOMANDĂRI ȘI STANDARDE PRIVITOARE LA REDACTAREA TEZEI DE DOCTORAT	11
STANDARDE PRIVIND STRUCTURA TEZEI DE DOCTORAT	11
STANDARDE PRIVIND FORMATAREA	13
STANDARDE PRIVIND CITAREA AUTORILOR	14
GLOSAR DE TERMENI SPECIFICI APARATULUI CRITIC	28
EVITAREA PLAGIATULUI	31
ASPECTE JURIDICE ȘI ETICE ALE PLAGIATULUI	31
CE ESTE PLAGIATUL?	32
TIPURI DE PLAGIAT	33
EVITAREA PLAGIATULUI	33
„NU AM VRUT SĂ PLAGIEZ!”	34
EXEMPLU	34
DATE DESPRE SUBIECȚII INTERVIEVAȚI	35
CONCLUZIE	36
ANEXE	37
THE RESEARCH PROCESS	53
QUALITATIVE OR QUANTITATIVE	78
THEORIES IN SCIENTIFIC RESEARCH	570
RESEARCH WRITING	1072

CONCEPEREA, STRUCTURA ȘI PLANIFICAREA UNEI TEZE DE DOCTORAT

Definirea, tematica și specificitatea unei teze de doctorat

Teza de doctorat este o lucrare academică care propune o abordare teoretică și/sau aplicativă bazată pe cercetări recente în domeniu în care autorul aduce o contribuție în domeniul respectiv prin extinderea cercetărilor pe o anumită tematică, prin critica modelelor teoretice existente și propunerea de noi modele și abordări în domeniul de cercetare ales, sau prin structurarea într-un mod original a liniilor de cercetare din domeniul respectiv.

Trăsătura distinctivă a unei teze de doctorat este constituită de investigarea științifică, argumentativă, a unei teme de cercetare precis precizate; ea nu este un demers jurnalistic sau literar, ci o abordare care utilizează metodologii specifice unui demers de cercetare – argumentarea alegerii temei de cercetare, sinteza pozițiilor teoretice cu privire la problema studiată, analiză critică a cercetărilor deja existente în domeniul respectiv, propunerea de noi teme sau ipoteze de cercetare, validarea experimentală, empirică a ipotezelor întrebărilor de cercetare, formularea de concluzii științifice etc. În acest sens, lectura articolelor de specialitate din reviste de cercetare specifice domeniului de cercetare are un dublu beneficiu: 1) reprezintă o importantă sursă informațională și 2) familiarizează cititorul cu abordarea și rigorile domeniului de cercetare.

În alegerea temei sunt importante/relevante următoarele aspecte:

- ce interese anterioare ați avut cu privire la tematica respectivă – teoretice sau practice;
- cât de familiar vă este, conceptual și faptic, domeniul în care doriți să realizați teza de doctorat;
- competențele de cercetare – ce metodologie stăpâniți și cum puteți opera cu aceasta;
- interesul și competența coordonatorului în domeniul respectiv;
- actualitatea temei respective.

Titlul tezei de doctorat

Titlul tezei de doctorat trebuie să informeze auditorul/publicul care este tema centrală a cercetării. Acesta este, în general, stabilit la începutul demersului de redactare a tezei împreună cu profesorul coordonator și poate fi schimbat pe parcurs, dacă redactarea tezei evoluează în direcții diferite sau aprofundează anumite aspecte descoperite pe parcurs ca fiind mai relevante.

Titlul se va referi la:

- tema specifică de cercetare și, eventual, ipoteza de cercetare;
- abordarea specifică teoretică sau metodologică;
- rezultatele sau impactul cercetării.

Greșeli posibile în formularea titlului:

- foarte general;
- imprecis;
- jurnalistic;
- foarte lung;
- metaforic.

Dimensiunea și structura tezei de doctorat

(I) Dimensiunea tezei de doctorat

În general, o *teză de doctorat are între 170 și 250 pagini*. Este important ca numărul minim de 170, să fie respectat, însă o teză de doctorat poate depăși 250 pagini (dacă tema se pretează sau dacă studenta/studentul - doctorand consideră că nu poate epuiza tema investigată în numărul de pagini recomandat).

Dimensiunea tezei de doctorat depinde de:

- cutumele domeniului în care se încadrează teza;
- aparatul teoretic necesar argumentării;

- tipul de metodologie de cercetare utilizat;
- volumul datelor colectate și analizate.

(II) Structura tezei de doctorat

Structura canonică a unei teze de doctorat este constituită din trei părți principale: *introducerea*, *corpul* și *concluziile* lucrării. Astfel, din totalul de 250 pagini ale tezei de doctorat, este recomandat ca între 10 și 15 pagini să fie dedicate *introducerii*, 5 și 10 pagini *concluziilor*, iar restul de 30, respectiv 50, de pagini să fie alocate fiecarui capitol care constituie corpul lucrării.

Această structură canonică trebuie detaliată pornind de la construirea unei scheme a lucrării care să cuprindă titlurile capitolelor și subcapitolelor. Schema lucrării a) facilitează delimitarea sferei de cercetare (astfel încât să poată fi evitate temele mult prea vaste care nu pot fi acoperite de o teză de doctorat) și b) reprezintă planul de lucru ale cărui întrebări de cercetare ghidează demersul științific ajutând la profilarea firului roșu al lucrării.

În continuare vom caracteriza structura canonică a unei teze de doctorat.

A) În *Introducere* trebuie specificate, în mod obligatoriu, următoarele elemente:

- 1) importanța temei de cercetare și motivația studentei/studentului–doctorand în alegerea temei,
- 2) întrebările de cercetare de la care pornește cercetarea,
- 3) metodologia utilizată,
- 4) structura pe capitole a tezei de doctorat și o scurtă descriere a acestora,
- 5) bibliografia și documentarea pe care se va construi cercetarea.

B) *Corpul* tezei de doctorat trebuie să fie constituit de elaborarea argumentelor subsumate întrebărilor de cercetare. Argumentele propuse trebuie să fie fundamentate metodologic, și, dacă este cazul, empiric. O trăsătură distinctivă a unui demers de cercetare cum este cel reprezentat de teza de doctorat constă în elaborarea cu rigoare științifică a argumentelor propuse. Argumentele pot susține două tipuri de rezultate:

- a) rezultate negative – a căror funcție este de a expune și critica deficiențele unei abordări sau ale unei teorii și
- b) rezultate pozitive – a căror funcție este de a propune abordări originale, de a teoretiza anumite aspecte eludate, insuficient sau inadecvat analizate.

C) În *Concluzii*, studenta/studentul-doctorand trebuie să indice care sunt rezultatele cercetării, care sunt răspunsurile la întrebările de cercetare propuse, ce limite au fost întâmpinate pe parcursul acestui demers de cercetare, dacă e cazul (de pildă, noutatea temei și lipsa unor cărți pe această temă) și modalitatea în care această cercetare deschide noi direcții de studiu care ar putea fi exploatate în viitor. De asemenea, studenta/studentul-doctorand trebuie să precizeze semnificația rezultatelor obținute din perspectivă teoretică și/sau practică în contextul cercetărilor din domeniu. Acolo unde este cazul, studenții-doctoranzi sunt încurajați să evalueze și prezinte impactul cercetării lor atât la nivel teoretic cât și practic.

Nota bene:

Exprimarea trebuie să dovedească însușirea unui limbaj de specialitate din domeniul în care se încadrează tema tezei de doctorat. Formulările adecvate sunt:

- demersul propus *analizează, tratează, descrie, explică etc.*;
- capitolul 1 *își propune prezentarea, analiza, observarea...etc.*;
- *datele au fost măsurate, extrase, verificate.*

Rolul coordonatorului

Pentru a surprinde mai bine rolul și limitele implicării coordonatorului în elaborarea tezei de doctorat vom defalca și sintetiza principalele atribuții și limite ale implicării coordonatorului:

Principalele atribuții ale coordonatorului constau în:

- indicarea abordării și structurii lucrării;
- analiza și validarea abordării și argumentării teoretice;
- analiza și validarea demersului metodologic;
- sprijin în selectarea bibliografiei relevante;
- sprijin în analiza și interpretarea datelor;
- feedback cu privire la progres;
- formularea de opinii critice;
- evaluarea lucrării pe parcurs și formularea de sugestii cu privire la progresul necesar pentru finalizarea acesteia și prezentarea publică.

*Coordonatorului **nu** îi revine responsabilitatea:*

- de a căuta și oferi studenților-doctoranzi bibliografia completă a tezei de doctorat (ci doar de a orienta studenta/studentul-doctorand în direcția unor lucrări relevante în domeniu). Studenții-doctoranzi trebuie să identifice, să inventarieze și să citească lucrări de specialitate, articole științifice corespunzătoare disciplinei în care se încadrează propriul demers, care pot fi găsite atât în biblioteci, cât și în bazele de date internaționale (ex: *J-Stor*, *Ebsco*). Menționăm că accesul la aceste baze de date internaționale se face fie la sala de lectură a Bibliotecii Facultății de Studii Europene, fie la sala Multimedia a Bibliotecii Central Universitare Lucian Blaga);
- de a oferi informații de ordin administrativ (studenta/studentul-doctorand are posibilitatea de a găsi aceste informații în Regulamentul Școlii doctorale sau de a le obține de la secretariat);
- de a sintetiza, în locul studenților-doctoranzi, partea teoretică a tezei de doctorat, ci doar de a dirija studenții-doctoranzi în acest sens;
- de a concepe, în locul studenților-doctoranzi, partea analitică a tezei de doctorat, ci doar de a coordona studenții-doctoranzi în acest sens. În esență, studenții-doctoranzi trebuie să aducă ceva nou în domeniul de studiu în care se încadrează lucrarea lor științifică, iar acest lucru este doar supervizat de către coordonator.

Conceperea și planificarea tezei de doctorat

Conceperea tezei de doctorat

Primul pas în conceperea tezei de doctorat constă în familiarizarea cu domeniul de cercetare și terminologia de specialitate. Preambulul cercetării este reprezentat de încadrarea temei în domeniu de cercetare (relații internaționale, studii europene, științe politice, economie, științe juridice, sociologie, antropologie, istorie etc) și o bună operare cu terminologia de specialitate. În funcție de aceste repere se va face și contactarea coordonatorului în vederea a) stabilirii titlului b) structurii generale a tezei c) planului/design-ului de cercetare. După această etapă, studentul/studenta-doctoranda și coordonatorul vor identifica principalele surse bibliografice corespunzătoare fiecărei componente a structurii generale a lucrării. Apoi, se va stabili de comun acord un calendar pentru parcurgerea acestei prime liste bibliografice. După

parcursarea primei liste bibliografice se trece la discutarea abordării teoretice și a părții practice (dacă este cazul), se va completa bibliografia relevantă a ambelor segmente și se va formula detaliat structura tezei de doctorat. În următoarea etapă, studentul/studenta-doctoranda va elabora și transmite coordonatorului sinteza părții teoretice a lucrării; această etapă e urmată de analiza și reformularea, pe baza sugestiilor coordonatorului, a părții teoretice a lucrării. Odată încheiate aceste etape, atenția se va concentra asupra elaborării studiului de caz, respectiv, realizării părții aplicative/experimentale a lucrării. În această etapă, se colectează și analizează datele relevante pentru studiul de caz sau cele obținute în urma aplicației implementate, după care se trece la interpretarea datelor în contextul teoretic prezentat în prima parte a lucrării, iar apoi formularea concluziilor. Colectarea, analiza, interpretarea și formularea concluziilor se face în deplină rigoare metodologică, iar ultima etapă se realizează prin stabilirea cu coordonatorul a aspectelor metodologice și a concluziilor formulate.

Etape în redactarea tezei de doctorat

Pentru eficientizarea demersului dumneavoastră vă propunem următorul algoritm de redactare a tezei de doctorat:

1. Selectarea și revizuirea bibliografiei

- utilizați variate modalități de căutare și selectare a bibliografiei – biblioteci, baze de date online, site-uri acreditate etc.;
- studiați diferite abordări metodologice specifice tematicii dvs.;
- verificați dacă aveți la dispoziție suficiente surse bibliografice – discutați selecția dvs. finală (pe baza listei autor/concept, abordare teoretică) cu profesorul coordonator.

2. Parcurgerea bibliografiei

- construiți o corespondență între abordările teoretice sintetizate și/sau angajate în lucrarea dumneavoastră și cercetătorii care le-au propus și utilizat;
- construiți o corespondență între conceptele teoretice utilizate și cercetătorii care le-au teoretizat, analizat, criticat;
- realizați o hartă conceptuală a lucrării pornind de la bibliografie.

3. Analiza și redactarea părții teoretice

- expuneți într-un mod nedistorsionat toate abordările teoretice angajate;
- expuneți proporțional toate abordările teoretice angajate;
- în prezentarea abordărilor teoretice nu omiteți expunerea criticilor și semnalarea deficiențelor acestora;
- încercați să exemplificați într-un mod original utilizarea abordărilor teoretice expuse.

4. Redactarea și formularea detaliată a abordării metodologice

- selectați împreună cu coordonatorul bibliografia pentru partea de metodologie și începeți redactarea acestei părți;
- prezentați un cadru comprehensiv asupra diferitelor modalități de abordare a studiului de caz sau a aplicației temei dvs. de cercetare;
- analizați diferitele abordări metodologice prezentând punctele forte și deficiențele fiecăreia;
- argumentați opțiunea metodologică făcută de dumneavoastră;
- discutați cu coordonatorul abordarea metodologică pentru care optați și revizuiți partea metodologică, dacă se impune.

5. Derularea cercetării

- pe baza planului de cercetare adoptat, colectați, analizați și interpretați datele;
- precizați limitele demersului metodologic sau problemele de derulare a acestuia.

6. Colectarea și prelucrarea datelor

- din opțiunile delimitate de cadrul metodologic asumat, selectați cele mai fezabile metode de colectare a datelor;
- culegeți datele respectând integral și cu strictețe protocoalele metodologice;
- verificați de cel puțin două ori acuratețea și rigurozitatea datelor colectate cu profesorul coordonator și începeți prelucrarea acestora.

7. Redactarea părții practice și analiza datelor

- folosind instrumentarul metodologic asumat utilizați datele colectate pentru a infirma sau confirma anumite ipoteze de cercetare;

- folosiți atât formulări textuale cât și imagini, grafice, tabele care să prezinte rezultatele relevante ale cercetării și care să susțină demersul dvs. de cercetare, respectiv, interpretarea pe care o prezentați.

8. Interpretarea datelor

- este una din cele mai importante părți ale demersului dvs. și cea care prezintă cel mai mare interes în evaluarea lucrării dvs.;
- validați interpretarea datelor cu profesorul coordonator;
- evitați interpretările prea generale sau prea vagi;
- formulați interpretările în relație cu temele/ipotezele de cercetare și obiectivele cercetării.

9. Expunerea concluziilor

- concluziile trebuie să ofere o imagine de ansamblu asupra a) semnificației demersului dvs. în contextul cercetărilor din același perimetru, b) limitelor și valorii adăugate a cercetării și c) direcțiilor de cercetare care pot fi explorate ulterior pornind de la lucrarea dvs.;
- concluziile trebuie să precizeze explicit care sunt răspunsurile rezultate în urma cercetării dvs. la întrebările de cercetare.

REGULI, RECOMANDĂRI ȘI STANDARDE PRIVITOARE LA REDACTAREA TEZEI DE DOCTORAT

Standarde privind structura tezei de doctorat

Coperta și Pagina de titlu – informațiile care trebuie să apară pe coperta și pagina de titlu a tezei de doctorat și tiparul acestora sunt prezentate în Anexele 1 și 2 ale prezentului Ghid.

Declarație standard – teza de doctorat va conține o declarație pe propria răspundere a absolventului, datată și *semnată olograf*, din care să rezulte că lucrarea îi aparține, nu a mai fost niciodată prezentată și nu este plagiată. Comisia va respinge, indiferent de momentul descoperirii, lucrările care conțin elemente plagate. Plagiatul înseamnă preluarea integrală sau parțială a unor texte, date sau idei fără referințe corespunzătoare la autorii acestora. Pentru detalii privind definirea plagiatului și regulile de evitare ale acestuia consultați capitolul corsepunzător din Ghid. Menționăm, de asemenea, două surse importante de documentare cu privire la plagiat și evitarea acestuia disponibile on-line la adresele:

<http://oxforddictionaries.com/definition/english/plagiarism>

<http://www.princeton.edu/writing/university/resources/WPAPlagiarism.pdf>

Cuprins – teza de doctorat va avea un cuprins care să conțină cel puțin titlurile tuturor capitolelor și subcapitolelor însoțite de numărul paginii la care începe fiecare dintre acestea.

Lista figurilor și lista tabelor – în cazul în care teza de doctorat conține figuri (imagini, grafice) și/sau tabele, acestea vor fi prezentate, imediat după cuprins, sub forma unor liste (separat pentru figuri și tabele) care conțin numele fiecărui element și numărul paginii la care se află acesta.

Introducere – aceasta va conține motivația alegerii temei, gradul de noutate al temei, obiectivele generale ale lucrării, metodologia folosită, un rezumat pe capitole al lucrării, precum și limitele/neaajunsurile lucrării (confidențialitatea datelor, rată mică de răspuns la chestionare/interviuri, lipsa accesului la unele surse bibliografice de referință etc.). De

asemenea, introducerea trebuie să precizeze care este relevanța lucrării, precum și originalitatea acesteia.

Tot în introducere, autorul trebuie să ofere detalii despre metodologia folosită, cu detalierea instrumentelor sau metodelor de cercetare. Totodată, trebuie prezentate ipotezele (dacă e vorba de o lucrare care folosește metode cantitative) și structura argumentelor dezvoltate. Tot în acest punct trebuie descrise sau precizate întrebările de cercetare care să susțină obiectivele cercetării. Atragem atenția asupra formulei obiectivelor și a întrebărilor de cercetare: dacă primele pot fi exprimate într-o manieră generală, întrebările de cercetare trebuie să fie precise și să delimiteze un domeniu îngust.

Ca regulă generală, sugerăm ca introducerea să ocupe 5-10% din textul integral al lucrării.

Corpul – teza de doctorat va conține între 3 și 6 capitole, numerotate crescător, fiecare putând avea (dacă vă ajută) o scurtă introducere, iar în partea finală, o secțiune de concluzii, care să sintetizeze informațiile și/sau rezultatele prezentate în cadrul aceluia capitol.

Capitolele trebuie astfel propuse încât să acopere cele două axe fundamentale ale unei teze de doctorat: partea teoretică și studiul de caz/aplicația.

Partea teoretică (întinsă pe parcursul unuia sau mai multe capitole, de regulă, unul singur) trebuie să conțină în mod obligatoriu o incursiune prin literatura de specialitate aferentă temei abordate. Pentru a emite opinii despre o anumită temă, este necesară parcurgerea, măcar parțială, a textelor publicate anterior în domeniul de cercetare vizat, precum și identificarea, dacă este cazul, a principalelor teorii/curente de gândire.

În capitolele de cercetare sau analiză propriu-zise trebuie evitat stilul eminent descrisiv. O teză de doctorat reprezintă punctul de vedere propriu al autorului, astfel că trebuie evitate preluările sau parafrazările. Nu în ultimul rând, trebuie avută în vedere originalitatea lucrării: chiar dacă textul aduce puține elemente noi, acestea trebuie evidențiate și prezentate cu prioritate.

Ultimele capitole conțin analiza unui studiu de caz sau prelucrarea unor date obținute prin cercetare, în cazul în care lucrarea operează cu date. Indiferent de formula de cercetare, autorul va pune în evidență ce anume aduce nou lucrarea față de studiile deja cunoscute.

Concluziile lucrării – în această parte a tezei de doctorat se regăsesc cele mai importante aserțiuni din lucrare, opinia calificată privind rezultatele obținute în lucrare, semnificația acestora precum și potențiale direcții viitoare de cercetare legate de tema abordată.

Bibliografia – aceasta este ultima parte a lucrării și va conține lista tuturor surselor de informație utilizate de către absolvent pentru redactarea lucrării de licență. Bibliografia trebuie

defalcată pe secțiuni, în funcție de statutul surselor consultate (surse primare, surse secundare, cărți, articole, rapoarte/studii etc.) De preferință ar trebui să conțină mai mult de 20 de titluri care, bineînțeles, trebuie să se regăsească citate pe parcursul lucrării.

Anexe (dacă este cazul) – acestea apar într-o secțiune separată, care nu se numerează. Fiecare anexă se va menționa cel puțin o dată în textul lucrării. Anexele se numerează crescător (Anexa 1, Anexa 2 etc.). Paginile care constituie Anexele nu se numerează, deoarece ele nu reprezintă produsul efortului intelectual al autorului, ci doar suport pentru documentare.

Nota bene:

- Introducerea și concluziile nu se numerează!
- Neconcordanța dintre notele de subsol și Bibliografie poate fi penalizată.
- Lucrările menționate în notele de subsol trebuie să se regăsească în Bibliografie.

Standarde privind formatarea

Pagină. Numărul minim de pagini este 170, în format A4. Se vor utiliza următoarele valori pentru marginile paginii (Page Setup -> Margins): stânga: 2,5 cm; dreapta: 2 cm; sus: 2 cm; jos: 2 cm. Numerotarea paginilor se face începând cu pagina de titlu, până la ultima pagină a lucrării, dar numărul paginii apare doar începând cu Introducerea. Numărul de pagină se inserează în subsolul paginii, centrat.

Paragraf. Textul va respecta o spațiere între rânduri de 1,5 linii (Format->Paragraph->Line spacing-> 1,5 lines). Textul din cadrul paragrafelor normale va fi aliniat între marginile din stânga și dreapta (*justified*). Primul rând al fiecărui paragraf va avea o indentare de 1,27 cm. Excepție fac titlurile capitolelor, care pot fi aliniate centrat sau la dreapta, precum și etichetele tabelelor și figurilor (a se vedea explicațiile de mai jos).

Font/caracter. Fontul utilizat pentru redactare va fi Times New Roman, cu dimensiunea de 12 puncte, utilizând diacriticele specifice limbii în care este redactată lucrarea (ă, ș, ț, î, â - pentru limba română). Folosirea unor expresii consacrate – în general expresii în latină – precum *a priori*, *in nuce*, *ad nauseam*, *in extenso* se face cu ajutorul fontului *italic*. De asemenea, font-ul *italic* se folosește pentru a sublinia importanța unor cuvinte, concepte, expresii etc. – la fel și mai puțin esteticul font underline – iar font-ul **bold** este folosit, în general, pentru scrierea capitolelor, subcapitolelor și secțiunilor lucrării. Combinarea aleatoare sau/și

excesivă a acestor moduri de a sublinia (ca să nu mai vorbim de folosirea lor concomitentă; de ex. **exces de subliniere**) subminează însuși demersul punerii în evidență. La fel și în cazul jocului cu tipul sau cu dimensiunea font-ului.

Tabele și Figuri (dacă e cazul). Tabelele se numerotează cu 2 cifre, prima reprezentând numărul capitolului, iar cea de a doua reprezentând numărul tabelului din capitolul respectiv. Fiecare tabel are număr și titlu, care se menționează deasupra tabelului, aliniat la marginea din dreapta. Dacă este cazul, sursa datelor se precizează sub tabel, aliniat între marginile din stânga și dreapta (*justified*), indicând în mod obligatoriu numele autorului(lor), lucrarea (cartea), editura, anul, pagina sau adresa de Internet completă. Figurile (aici sunt incluse imagini, grafice, capturi de ecran) se numerotează cu 2 cifre, prima reprezentând numărul capitolului, iar cea de a doua fiind numărul figurii din capitolul respectiv; fiecare figură are număr și titlu, care se menționează deasupra figurii, centrat; dacă este cazul, sursa figurii se indică sub figură, *justified*, indicând numele autorului(lor), lucrarea (cartea), editura, anul, pagina sau adresa de Internet completă.

Standarde privind citarea autorilor

Ocaziile citării

În cazul parafrazării, trebuie semnalată sursa ideilor sau argumentelor prezentate. În cazul reproducerii, pe lângă semnalarea sursei, trebuie să folosiți ghilimelele, pentru a preciza preluările literale.

În concluzie, sursa trebuie precizată în orice situație în care o idee sau o structură argumentativă a unui autor este preluată, fie general-descriptiv, fie literal.

Stilul citării

Propunem stilul „clasic” de citare, cu note de subsol, după următorul model (numele, titlurile și numărul paginilor, deși sunt reale, au fost alese în mod aleator).

Dacă specificul lucrării elaborate și surselor citate vă permit, puteți utiliza stilul „Harvard” de citare, prezentat în Anexe, **cu acordul cadrului didactic coordonator**.

1. Cărți

A. Cărți tipărite

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, *Titlul cărții cu caractere italice*, Ediția cărții (dacă este cazul), Orașul: Editura, anul, pagina/paginile.

Exemple:

(i1) *un autor:* Pierre Bourdieu, *Rațiuni practice. O teorie a acțiunii*, București: Editura Meridiane, 1999, p. 93.

(i2) *doi autori:* Ian Chiswell și Wilfrid Hodges, *Mathematical Logic*, Oxford: Oxford University Press, 2007, pp. 171-173 (în cazul în care informația se extinde pe mai multe pagini, folosiți pp. și dați intervalul paginilor).

(i3) *trei autori:* George S. Boolos, John P. Burgess și Richard C. Jeffrey, *Computability and Logic*, ediția a IV-a, New York: Cambridge University Press, 2002, p. 127.

(i4) *patru sau mai mulți autori:* Edward O. Laumann *et al.*, *The social organization of sexuality: Sexual practices in the United States*, Chicago: University of Chicago Press, 1994, p. 55.

Nota bene:

a) În elaborarea notelor de subsol, între fragmentele de informații (autor, titlu etc.) se intercalează virgule, nu puncte, cu excepția menționării orașului și editurii, între care se intercalează simbolul două puncte (:)

b) În cazul în care cartea are patru sau mai mult de patru autori citați în text doar primul autor urmat de *et al.*

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, *Titlul cărții cu caractere italice*, Oraș: Editura, anul.

Exemple:

(ii1) *un autor:* Bourdieu, Pierre, *Rațiuni practice. O teorie a acțiunii*, București: Editura Meridiane, 1999.

(ii2) *doi autori:* Chiswell, Ian și Hodges, Wilfrid, *Mathematical Logic*, Oxford: Oxford University Press, 2007.

(ii3) *trei autori:* Barth, James; Caprio, Gerard și Levine, Ross, *Rethinking Bank Regulation: Till Angels Govern*, New York: Cambridge University Press, 2006.

(ii4) *patru sau mai mulți autori*: Laumann, Edward O. *et al.*, *The social organization of sexuality: Sexual practices in the United States*, Chicago: University of Chicago Press, 1994.

Nota bene:

- a) Pentru ordonarea alfabetică a bibliografiei cităm mai întâi numele apoi prenumele fiecărui autor despărțite de virgulă.
- b) Citarea cărților cu mai mulți autori se face în ordinea în care aceștia apar în carte.
- c) În cazul unei cărți scrise de trei autori, la bibliografie specificăm numele și prenumele primului autor, apoi inserăm simbolul pentru punct și virgulă (;) după care specificăm numele și prenumele celui de-al doilea autor urmat de conjuncția *și*, apoi specificăm numele și prenumele celui de-al treilea autor.
- d) Toate sursele folosite în lucrare trebuie indicate și în bibliografie.
- e) În bibliografie nu reluăm paginile citate.
- f) Observați că la finalul fiecărei citări/referințe bibliografice apare un punct, la fel ca în cazul oricărei propoziții.

B. Cărți în format electronic

B1. Cărți accesibile online

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, *Titlul cărții cu caractere italice*, Orașul: Editura, anul, pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

Adrian Ludușan, *Logică Matematică*, Cluj-Napoca: Presa Universitară Clujeană, 2013, p. 27, <http://www.editura.ubbcluj.ro/bd/ebooks/pdf/1717.pdf>, accesat în 14.06.2015.

Nota bene:

În cazul în care cartea consultată are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, *Titlul cărții cu caractere italice*, Orașul: Editura, anul, adresa URL (link-ul).

Exemplu: Ludușan, Adrian, *Logică Matematică*, Cluj-Napoca: Presa Universitară Clujeană, 2013, <http://www.editura.ubbcluj.ro/bd/ebooks/pdf/1717.pdf>.

B2. Cărți în format Kindle, EPUB, sau un format similar.

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, *Titlul cărții cu caractere italice*, Orașul: Editura, anul, pagina/paginile, Ediția (urmată de) Tipul formatului (Kindle, EPUB).

Exemplu:

George S. Boolos, John P. Burgess și Richard C. Jeffrey, *Computability and Logic*, ediția a IV-a, New York: Cambridge University Press, 2002, p. 127, ediția Kindle.

Nota bene:

În cazul în care cartea consultată are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, *Titlul cărții cu caractere italice*, Orașul: Editura, anul, Ediția (urmată de)Tipul formatului (Kindle, EPUB).

Exemplu:

Boolos, George S.; Burgess, John P. și Jeffrey, Richard C., *Computability and Logic*, ediția a IV-a, New York: Cambridge University Press, 2002, ediția Kindle.

2. Capitole sau articole dintr-un volum colectiv (tipărite sau în format electronic)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului încadrat de ghilimele” în Prenumele și Numele editorului/lor (ed.)/(eds.) sau coordonatorului/lor (coord.), *Titlul volumului cu caractere italice*, Orașul: Editura, anul, pagina/paginile.

Exemple:

(i1) *volum colectiv cu un editor:* Mark Steiner, „Mathematics – Application and Applicability” în Stewart Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford: Oxford University Press, 2005, p. 627.

(i2) *volum colectiv cu un coordonator:* Cornelia Mureșan, „Analiza logliniară” în Traian Rotariu (coord.), *Metode statistice aplicate în științele sociale*, Iași: Polirom, 2006, pp. 229-230.

(i3) *volum colectiv cu mai mulți editori*: Boris Groys, „Comunismul privit din afară” în Adrian T. Sîrbu și Alexandru Polgar (eds.), *Genealogii ale postcomunismului*, Cluj: Idea Design & Print, 2009, p. 48.

Nota bene:

- a) În cazul în care cartea consultată are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).
- b) În cazul în care cartea consultată este în format electronic, aplicați indicațiile de la punctul 1B.

- (ii) **Bibliografie**: Numele, Prenumele autorului/lor, „Titlul articolului încadrat de ghilimele”, în Prenumele și Numele editorului/lor (ed.)/(eds.), *Titlul cărții cu caractere italice*, Orașul: Editura, anul, paginile între care este publicat capitolul/articolul.

Exemple:

(ii1) *volum colectiv cu un editor*: Steiner, Mark, „Mathematics – Application and Applicability” în Shapiro, Stewart (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, Oxford: Oxford University Press, 2005, pp. 624-649.

(ii2) *volum colectiv cu un coordonator*: Mureșan, Cornelia, „Analiza logliniară” în Traian Rotariu (coord.), *Metode statistice aplicate în științele sociale*, Iași: Polirom, 2006, pp. 228-253.

(ii3) *volum colectiv cu mai mulți editori*: Groys, Boris, „Comunismul privit din afară” în Sîrbu, Adrian T. și Polgar, Alexandru (eds.), *Genealogii ale postcomunismului*, Cluj: Editura Idea Design & Print, 2009, pp. 47-58.

Nota bene:

- a) În cazul în care cartea consultată are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).
- b) În cazul în care cartea consultată este în format electronic, aplicați indicațiile de la punctul 1B.

3. Prefața, cuvânt înainte, introducere, studiu introductiv sau altă parte similară dintr-o carte/volum (tipărite sau în format electronic)

- (i) **Notă de subsol**: Prenumele și Numele autorului/lor, „Prefața/Cuvânt înainte/Introducere/Studiu introductiv etc.” în Prenumele și Numele autorului/lor,

Editorului/lor (ed.)/(eds.) sau coordonatorului/lor (coord.), *Titlul cărții cu caractere italice*, Orașul: Editura, anul, pagina/paginile.

Example:

Adrian Ludușan și Bogdan Dicher, „Foreword” în Adrian Ludușan, Bogdan Dicher (eds.) *Philosophy of Pragmatism (II) Varieties of Pragmatism. Classical Tradition and Contemporary Developments*, Cluj-Napoca: EFES, 2009, p. 5.

Mircea Dumitru, „Cuvânt înainte” în Willard van Orman Quine și Joseph Ullian, *Țesătura opiniilor*, Pitești: Paralela 45, 2007, p. 9.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Prefața/Cuvânt înainte/Introducere/Studiu introductiv etc” în Numele, Prenumele autorului/lor sau Editorului/lor (ed.)/(eds.), *Titlul cărții cu caractere italice*, Orașul: Editura, anul, paginile între care este publicată partea citată.

Example:

Ludușan, Adrian și Dicher, Bogdan, „Foreword” în Ludușan, Adrian și Dicher, Bogdan (eds.), *Philosophy of Pragmatism (II) Varieties of Pragmatism. Classical Tradition and Contemporary Developments*, Cluj-Napoca: EFES, 2009, pp. i-xiii.

Dumitru, Mircea, „Cuvânt înainte” în Quine, Willard van Orman și Ullian, Joseph, *Țesătura opiniilor*, Pitești: Paralela 45, 2007, pp. 5-15.

Nota bene:

- a) În cazul în care cartea consultată are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).
- b) În cazul în care cartea consultată este în format electronic, aplicați indicațiile de la punctul 1B.

4. Recenzii de cărți/volume editate, coordonate (tipărite sau în format electronic)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul recenziei încadrat cu ghilimele” (dacă este cazul), recenzie la *Titlul cărții/volumului cu caractere italice* de Prenumele și Numele autorului/lor sau Editorului/lor (ed.)/(eds.), în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul, anul, Secțiunea (Recenzii, Book Review), pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

Christopher Pincock, recenzie la *The Applicability of Mathematics in Science: Indispensability and Ontology* de Sorin Bangu, în *Philosophia Mathematica*, Vol. 23, No. 3, 2014, Critical Studies/Book Reviews, p. 405.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, Titlul recenziei încadrat cu ghilimele” (dacă este cazul), recenzie la *Titlul cărții/volumului cu caractere italice* de Prenumele și Numele autorului/lor sau Editorului/lor (ed.)/(eds.), în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul, anul, Secțiunea (Recenzii, Book Review), adresa URL (link-ul), data accesării.

Exemplu:

Pincock, Christopher, recenzie la *The Applicability of Mathematics in Science: Indispensability and Ontology* de Bangu, Sorin, în *Philosophia Mathematica*, Vol. 23, No. 3, 2014, Critical Studies/Book Reviews.

Nota bene:

- a) În cazul în care cartea/volumul consultat are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).

5. Articole incluse în reviste

A. Articole incluse în reviste tipărite

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului încadrat de ghilimele”, în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, pagina/paginile.

Exemplu:

David Theo Goldberg, „Racism and Rationality: The Need for a New Critique” în *Philosophy of the Social Sciences*, Vol. 20, No. 3, September 1990, p. 344.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, paginile între care este publicat articolul.

Exemplu:

Goldberg, David Theo, „Racism and Rationality: The Need for a New Critique” în *Philosophy of the Social Sciences*, Vol. 20, No. 3, September, 1990, pp. 317-350.

Nota bene:

În cazul în care articolul consultat are mai mulți autori, aplicați indicațiile de la punctul 1A(i).

B Articole incluse în reviste electronice**B1. Reviste pentru care este disponibil un DOI (Digital Object Identifier)**

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, pagina/paginile, DOI, data accesării.

Exemplu:

Reuben Goodstein, „On the restricted ordinal theorem” în *Journal of Symbolic Logic*, Vol. 9, No. 2, 1944, p. 35, doi: [10.2307/2268019](https://doi.org/10.2307/2268019), accesat în 14.02.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, paginile între care este publicat articolul, DOI.

Exemplu:

Goodstein, Reuben, „On the restricted ordinal theorem” în *Journal of Symbolic Logic*, Vol. 9, No. 2, 1944, pp. 33-41, doi: [10.2307/2268019](https://doi.org/10.2307/2268019).

Nota bene:

În cazul în care articolul consultat are mai mulți autori, aplicați indicațiile de la punctul 1A(i).

B2. Reviste pentru care nu este disponibil un DOI – specificăm URL-ul (link-ul)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

Hilary Putnam, „Nonstandard Models and Kripke's Proof of the Gödel Theorem” în *Notre Dame Journal of Formal Logic*, Vol. 41, No. 1, 2000, p. 55,

https://projecteuclid.org/download/pdfview_1/euclid.ndjfl/1027953483, accesat în 21.10.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, Trimestrul (dacă este cazul), anul, paginile între care este publicat articolul, adresa URL (link-ul), data accesării.

Exemplu:

Putnam, Hilary, „Nonstandard Models and Kripke's Proof of the Gödel Theorem” în *Notre Dame Journal of Formal Logic*, Vol. 41, No. 1, 2000, pp. 53-58, https://projecteuclid.org/download/pdfview_1/euclid.ndjfl/1027953483.

Nota bene:

În cazul în care articolul consultat are mai mulți autori, aplicați indicațiile de la punctul 1A(i).

6. Articole în ziare, cotidiene, magazine (tipărite sau în format electronic)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, data, pagina/paginile (dacă este cazul), adresa URL (link-ul), data accesării.

Exemplu:

articol în ziar: Neil MacFarquhar, „For Russia, Links Between Caucasus and ISIS Provoke Anxiety” în *New York Times*, 20 noiembrie 2015, <http://www.nytimes.com/pages/world/europe/index.html?action=click&contentCollection=Europe&module=Kicker®ion=Header&pgtype=article>, accesat în 20.11.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul articolului încadrat de ghilimele” în *Titlul revistei cu caractere italice*, Volumul, Numărul, data, paginile între care este publicat articolul (dacă este cazul), adresa URL (link-ul).

Exemplu:

articol în ziar: MacFarquhar, Neil, „For Russia, Links Between Caucasus and ISIS Provoke Anxiety” în *New York Times*, 20 noiembrie 2015,

<http://www.nytimes.com/pages/world/europe/index.html?action=click&contentCollection=Europe&module=Kicker®ion=Header&pgtype=article>.

Nota bene:

- a) În cazul în care articolul consultat are mai mulți autori sau mai multe ediții, aplicați indicațiile de la punctul 1A(i).
- b) În cazul în care lipsesc informații cu privire la unii itemi trebuie citate toate referințele ce pot fi găsite cu privire la textul citat: autorul (dacă apare), numele articolului, tema/rubrica din care face parte (dacă este cazul), numărul sau data publicării pe site.
- c) În cazul în care articolul nu are paginație menționați secțiunea cea mai apropiată din care face parte citatul.

7. Articole/Documente preluate de pe platforme online sau pagini Web.

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul articolului/documentului încadrat de ghilimele”, *Numele platformei/website-ului cu caractere italice*, data postării (dacă este accesibilă), Secțiunea (dacă există), pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

(i1) *articol preluat de pe o platformă online:* Vasile Ernu, „Totul rămîne pe vechi sau cauzele înfrîngerii lui Saakashvili”, *CriticAtac*, 3 octombrie 2012,
<http://www.criticatac.ro/19333/totul-rmine-pe-vechi-sau-cauzele-infringerii-lui-saakavili/>, accesat în 03.10.2012.

(i2) *document preluat de pe o pagină web:* Peter Smith, „Tennenbaum's Theorem”, *Logic Matters*, 28 februarie 2014, p. 4,
http://www.logicmatters.net/resources/pdfs/tennenbaum_new.pdf, accesat în 14.02.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul articolului/documentului încadrat de ghilimele”, *Numele platformei/website-ului cu caractere italice*, data postării (dacă este accesibilă), paginile între care este încadrat articolul (dacă este cazul), adresa URL (link-ul).

Exemplu:

(i1) *articol preluat de pe o platformă online*: Ernu, Vasile, „Totul rămîne pe vechi sau cauzele înfrîngerii lui Saakașvili”, *CriticAtac*, 3 octombrie 2012, <http://www.criticatac.ro/19333/totul-rmine-pe-vechi-sau-cauzele-infringerii-lui-saakavili/>.

(i2) *document preluat de pe o pagină web*: Smith, Peter, „Tennenbaum's Theorem”, *Logic Matters*, 28 februarie 2014, http://www.logicmatters.net/resources/pdfs/tennenbaum_new.pdf.

Nota bene:

- a) În cazul în care lipsesc informații cu privire la unii itemi trebuie citate toate referințele ce pot fi găsite cu privire la textul citat.
- b) În cazul în care documentul nu are paginație menționați secțiunea cea mai apropiată din care face parte citatul.
- c) Sursele web consultate se trec separat, într-un spațiu special alocat acestora.

8. Teze de doctorat, lucrări de disertație sau licență (tipărite sau în format electronic)

- (i) **Notă de subsol**: Prenumele și Numele autorului, *Titlul tezei, disertației sau licenței cu caractere italice*, Teză de doctorat/Lucrare de disertație/Lucrare de licență, Instituția de învățământ în cadrul căreia a fost susținută lucrarea, Orașul, anul, pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

Adrian Ludușan, *Teorii ale referinței*, Teză de doctorat, Universitatea Babeș-Bolyai, Cluj-Napoca, 2013, p. 45.

- (ii) **Bibliografie**: Numele, Prenumele autorului, *Titlul tezei, disertației sau licenței cu caractere italice*, Teză de doctorat/Lucrare de disertație/Lucrare de licență, Instituția de învățământ, Orașul, anul.

Exemplu:

Ludușan, Adrian, *Teorii ale referinței*, Teză de doctorat, Universitatea Babeș-Bolyai, Cluj-Napoca, 2012.

9. Documente oficiale (tipărite sau în format electronic – accesibile pe website)

- (i) **Notă de subsol:** *Titlul documentului cu caractere italice*, Volumul, Numărul, Trimestrul, Secțiunea, Orașul: Numele instituției care a elaborat documentul, anul, pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

Standard Eurobarometer, No. 81, Spring 2014, Public Opinion in the European Union, Bruxelles: European Commision, 2014, p. 55,
http://ec.europa.eu/public_opinion/archives/eb/eb81/eb81_publ_en.pdf, accesat în 21.10.2014.

- (ii) **Bibliografie:** Numele instituției care a elaborat documentul, *Titlul documentului cu caractere italice*, Volumul, Numărul, Trimestrul, Secțiunea, Orașul, anul, adresa URL (link-ul).

Exemplu:

European Commision, *Standard Eurobarometer*, No. 81, Spring 2014, Public Opinion in the European Union, Bruxelles, 2014,
http://ec.europa.eu/public_opinion/archives/eb/eb81/eb81_publ_en.pdf.

Nota bene:

- a) În cazul în care lipsesc informații cu privire la unii itemi trebuie citate toate referințele ce pot fi găsite cu privire la textul citat.
- b) În cazul în care documentul nu are paginație menționați secțiunea cea mai apropiată din care face parte citatul.

10. Blog – uri

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul intrării pe blog (postării) încadrat de ghilimele”, *Numele blog-ului cu caractere italice*, data postării, Secțiunea, adresa URL (link-ul), data accesării.

Exemplu:

Timothy Gowers, „Taylor’s theorem with the Lagrange form of the remainder”, *Gower’s Weblog*, 11 februarie 2014, Secțiunea: Taylor’s theorem with the Peano form of the remainder, <https://gowers.wordpress.com/2015/11/10/interesting-times-in-academic-publishing/#more-6003>, accesat în 14.02.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, „Titlul intrării pe blog (postării) încadrat de ghilimele”, *Numele blog-ului cu caractere italice*, data postării, adresa URL (link-ul).

Exemplu:

Gowers, Timothy, „Taylor’s theorem with the Lagrange form of the remainder”, *Gower’s Weblog*, 11 februarie 2014, <https://gowers.wordpress.com/2015/11/10/interesting-times-in-academic-publishing/#more-6003>.

Nota bene:

- a) În cazul în care lipsesc informații cu privire la unii itemi trebuie citate toate referințele ce pot fi găsite cu privire la textul citat.
 b) În cazul în care documentul nu are paginatie menționați secțiunea cea mai apropiată din care face parte citatul.
 c) *Nota bene:*

Sursele web consultate se trec separat, într-un spațiu special alocat acestora.

11. Lucrări prezentate în cadrul unor manifestări științifice (conferințe, congrese etc)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, „Titlul lucrării încadrat în ghilimele”, lucrare prezentată în cadrul manifestării științifice (conferință, congres) *Denumirea manifestării științifice cu caractere italice*, data manifestării, Instituția organizatoare.

Exemplu:

Adrian Ludușan, „On the significance of categoricity arguments”, lucrare prezentată în cadrul congresului *Congress of Logic, Methodology, and Philosophy of Science*, 3-8 august 2015, Universitatea din Helsinki.

- (ii) **Bibliografie:** Prenumele și Numele autorului/lor, „Titlul lucrării încadrat în ghilimele”, lucrare prezentată în cadrul manifestării științifice (conferință, congres) *Denumirea manifestării științifice cu caractere italice*, data manifestării, Instituția organizatoare.

Exemplu:

Ludușan, Adrian, „On the significance of categoricity arguments”, lucrare prezentată în cadrul congresului *Congress of Logic, Methodology, and Philosophy of Science*, 3-8 august 2015, Universitatea din Helsinki.

12. Lucrări nepublicate (manuscrise, draft-uri – în format fizic sau electronic)

- (i) **Notă de subsol:** Prenumele și Numele autorului/lor, *Titlul manuscrisului/draft-ului cu caractere italice* (dacă manuscrisul/draft-ul se prezintă sub forma unei cărți)/„Titlul manuscrisul/draft-ul încadrat în ghilimele”, (dacă manuscrisul/draft-ului se prezintă sub forma unui articol), manuscris/draft, data, pagina/paginile, adresa URL (link-ul), data accesării.

Exemplu:

(i1) *manuscris (neaccesibil online) sub formă de carte:* Shaughan Lavine, *Skolem was wrong*, manuscris, 1999, p. 21.

(i2) *draft (sub formă) de articol accesibil online:* Solomon Feferman, „Is the Continuum Hypothesis a definite mathematical problem?”, draft, 18.9.2011, p. 8, http://logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf, accesat în 14.02.2015.

- (ii) **Bibliografie:** Numele, Prenumele autorului/lor, *Titlul manuscrisului/draft-ului cu caractere italice* (dacă manuscrisul/draft-ul se prezintă sub forma unei cărți)/„Titlul manuscrisul/draft-ul încadrat în ghilimele”, (dacă manuscrisul/draft-ului se prezintă sub forma unui articol), manuscris/draft, data, adresa URL (link-ul).

Exemplu:

(i1) *manuscris (neaccesibil online) sub formă de carte:* Lavine, Shaughan, *Skolem was wrong*, manuscris, 1999.

(i2) *draft (sub formă) de articol accesibil online:* Feferman, Solomon, „Is the Continuum Hypothesis a definite mathematical problem?”, draft, 18.9.2011, http://logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf.

Glosar de termeni specifici aparatului critic

apud se folosește în cazul în care *citarea este preluată dintr-o altă sursă* decât cea orginară (din necunoașterea limbii în care a fost scrisă lucrarea, a comoditatății, a sărăciei bibliotecii etc.).

Exemplu:

ⁿKenneth Arrow, „The Principle of Rationality in Collective Decisions” în *Collected Papers of Kenneth J. Arrow*, Vol. 1, *Social Choice and Justice*, Cambridge MA: Belknap Press, 1984, p. 51, *apud* Amartya Sen, *Rationality and Freedom*, Cambridge, MA: Belknap Press, 2002, p. 328.

Contra se folosește pentru a specifica referințele la o poziție contrară celei exprimate.

Cf. sau *comp.* se utilizează pentru a specifica referințele la alte poziții (ușor diferite) pentru compararea sau confruntarea poziției exprimate.

[f.a.] se folosește în cazul în care anul publicării e necunoscut.

[f.e.] se folosește în cazul în care editura e necunoscută.

[f.l.] se folosește în cazul în care localitatea în care a apărut lucrarea e necunoscută.

Ibidem (aceeași lucrare) se folosește dacă vă referiți la lucrarea menționată în citarea imediat anterioară.

Exemplu:

ⁿIan Chiswell și Wilfrid Hodges, *Mathematical Logic*, Oxford: Oxford University Press, 2007, p. 171.

ⁿ⁺¹*Ibidem*, p. 173.

Idem (același autor) se folosește dacă citați succesiv două lucrări ale aceluiași autor.

Exemplu:

ⁿCharles Parsons, *Mathematical Thought and its Objects*, Cambridge: Cambridge University Press, 2008, pp. 17-19.

ⁿ⁺¹*Idem*, „The uniqueness of the natural numbers”, *Iyyun*, Vol. 39, 1990, p. 15.

Infra/vezi mai jos se folosesc pentru a preciza *pagina sau nota de subsol ulterioară unde se găsesc detalii* cu privire la tema sau ideea discutată în acest loc *sau sunt specificate referințele* la o lucrare menționată în acest punct.

op. cit. se folosește dacă între prima citare (completă) a lucrării unui autor și citările ulterioare se intercalează citarea unor lucrări ale unor autori diferiți .

Exemplu:

ⁿIan Chiswell și Wilfrid Hodges, *Mathematical Logic*, Oxford: Oxford University Press, 2007, p. 171.

ⁿ⁺¹Shaughan Lavine, *Skolem was wrong*, manuscris, 2001, p. 151.

ⁿ⁺² Chiswell și Hodges, *op. cit.*, p. 174.

passim/et passim (în diferite locuri) se folosesc pentru a preciza că *ideea menționată în lucrare se regăsește (difuz) pe mai multe pagini sau paragrafe* din proximitatea citării.

seq/et seq/sq/qq (și în următoarele) se folosesc pentru a preciza că *ideea menționată în lucrare este dezvoltată pe următoarele pagini, paragrafe* ale lucrării citate.

[*sic*] sau [*sic!*] se folosește pentru a indica *greșeli gramaticale, contradicții, inconsecvențe evidente* în cadrul citatului.

Supra/vezi mai sus se folosesc pentru a preciza *pagina sau nota de subsol anterioară unde se găsesc detalii* cu privire la tema sau ideea discutată în acest loc *sau sunt specificate referințele* la o lucrare menționată în acest punct.

Vezi și se folosește pentru a specifica *referințe la lucrări în care este expus sau argumentat un punct de vedere similar* cu cel prezentat.

*** se folosește dacă *autorul nu este specificat, este anonim sau necunoscut*.

Exemplu:

*** *Îndreptar ortografic, ortoepic și de punctuație*, Ediția a V-a, București: Univers Enciclopedic, 1995.

Nota bene:

- a) Formula *loc. cit.* se poate folosi în locul formulei *ibidem* sau formulei *op. cit.* dacă cităm același autor, aceeași carte, aceeași pagină. De remarcat, însă, e tendința de a înlocui formula *loc. cit.* cu *ibidem*.
- b) Observați omiterea prenumelui în cazul folosirii formulei *op. cit.*
- c) Citarea indirectă a unei surse, prin *apud*, presupune consemnarea în bibliografie doar a sursei secundare, indirecte, nu a sursei originare, neconsultate.
- d) După *Cf.* nu puneți virgulă.
- e) [*sic*] sau [*sic!*] se folosește în cadrul citatului, nu în specificațiile referențiale din nota de subsol.

EVITAREA PLAGIATULUI

Aspecte juridice și etice ale plagiatului

Plagiatul se referă la copierea unor idei, a unui raționament, a unui material vizual de tip fotografie, a unor pasaje de text, sau copierea integrală a unei opere de creație intelectuală fără a preciza sursa informației.

În țara noastră, Legea nr. 8/1996 consacrată protejării dreptului de autor nu prevede noțiunea de plagiat, întrucât plagiatul este mai degrabă un termen academic și mai puțin unul juridic. Corespondentul juridic pentru plagiat este „reproducerea fără drept” sau „încălcarea dreptului de autor”; altfel spus, plagiatul este sancționat în legislația națională sub o altă denumire juridică.

Plagiatul este sancționat și de către Universitatea Babeș-Bolyai. Codul de etică al Universității Babeș-Bolyai (Art. 22) clasifică încălcările eticii în domeniul cercetării astfel:

- a) plagiatul;
- b) omisiunea recunoașterii, fie prin menționare ca autor al unei opere, fie prin indicarea sursei, a contribuției unor terți la elaborarea unei opere;
- c) obligarea autorilor unei opere de a menționa ca autori și persoane care nu au participat la elaborarea acesteia;
- d) menționarea ca autori ai unei lucrări a unor persoane care nu au contribuit semnificativ la elaborarea acesteia¹.

La rândul ei, Facultatea de Studii Europene consideră esențială respectarea cadrului legal și a normelor deontologice prin acordarea creditului științific în scrierea tezei de doctorat. În caz contrar, dacă nu se specifică sursa informațiilor și se însușesc parțial sau în totalitate ideile ori cuvintele altora, se comite un furt intelectual. **Astfel, coordonatorul lucrării și comisia de specialitate vor respinge teza de doctorat în cazul în care se constată că lucrarea conține elemente plagiatare. În consecință, plagiatul este o problemă extrem de severă, chiar dacă nu**

¹ *Codul de etică și deontologie profesională al Universității Babeș-Bolyai*, Cluj-Napoca: Universitatea Babeș-Bolyai, 2013, p. 6, http://www.ubbcluj.ro/ro/despre/publice/files/Codul_Etic_al_UBB.pdf, accesat în 12.06.2015.

este comisă intenționat, ci din pură neatenție, având consecințe dramatice asupra studentului-doctorand.

Ce este plagiatul?

Dicționarul explicativ al limbii române definește plagiatul în următorul mod:

Acțiunea de a plagia; plagiarea = operă literară, artistică sau științifică a altcuiva însușită (integral sau parțial) și prezentată drept creație personală².

Aceasta înseamnă, de fapt, (1) a asuma ideile sau enunțurile altcuiva ca fiind proprii; (2) a le utiliza fără creditarea sursei; (3) a pretinde într-o lucrare că o idee este nouă, chiar dacă a fost derivată dintr-o altă sursă. Astfel, plagiatul poate fi considerat un act de fraudă din două motive: mai întâi, deoarece implică nerecunoașterea muncii depuse de o altă persoană, și, mai apoi, pentru că nu recunoaște importanța efortului depus de persoana care a fost furată.

Întrebarea formulată adesea de către studenți-doctoranzi este următoarea: **cum pot fi furate enunțurile și ideile altora?**

Răspunsul este destul de simplu. Ideile, interpretările date unor fenomene sau formulările originale sunt considerate proprietăți intelectuale și sunt protejate de legea dreptului de autor, tot așa cum sunt protejate și invențiile, brevetele etc. Aproape toate rezultatele obținute prin cercetare cad sub protecția dreptului de autor, atât timp cât acestea sunt înregistrate într-un fel (cum ar fi prin publicarea într-o carte, articol sau expunerea lor publică). Așadar, plagiatul înseamnă a copia enunțurile sau ideile unui autor și a pretinde că sunt proprii, a nu trimite la sursa de unde provine o idee, a omite ghilimelele, a furniza informații incorecte cu privire la sursa citată, a modifica succesiunea ideilor și a pretinde originalitatea lor, dar și copierea structurii unei lucrări. **Însă cazurile de plagiat pot fi evitate citând sursele de unde au fost culese informațiile**, adică, pur și simplu precizând ce materiale au fost folosite.

² *Dicționarul explicativ al limbii române*, varianta on-line: <http://dexonline.ro/definitie/plagiat>, accesat în 12.06.2015.

Tipuri de plagiat

Literatura de specialitate redă o varietate extinsă de tipuri de plagiat, însă mai toate cad de acord asupra unor categorii principale:

1. **Copiere integrală** – se prezintă capitole sau pasaje întregi dintr-o altă lucrare, după tehnica *copiere-lipire* (celebrul *copy-paste*), pretinzând ca acestea ar fi creație proprie. În această categorie intră și **lucrările cumpărate** de la terți sau primite **cadou** de la colegi.
2. **Copiere parțială** – copierea câtorva rânduri, a unor paragrafe sau părți semnificative (lungi de câteva pagini) dintr-o lucrare.
3. **Copiere prin parafrizare** – se preia structura, linia de argumentare, exemplele și alte elemente de conținut ale sursei, dar se modifică formulările, succesiunea paragrafelor sau alte elemente pentru a face mai dificilă identificarea autorului.
4. **Auto-plagiere** – se preia integral sau masiv dintr-o lucrare proprie, anterioară, inclusiv propria lucrare de licență sau disertație³.

Evitarea plagiatului

Evitarea plagiatului presupune urmarea câtorva pași în redactarea propriilor lucrări:

1. În primul rând, autorul trebuie să conceapă o lucrare **originală** și să ofere un plus de cunoaștere în domeniu.
2. Mai apoi, el trebuie să precizeze în aparatul critic sursele utilizate, în principal pentru a oferi **credibilitate** poziției susținute de autorul tezei de doctorat.
3. Specificarea precisă a autorului și a sursei prin **citare** (conform standardelor de citare).
4. Evitarea formulărilor de tipul „este cunoscut faptul”, „alți cercetători au arătat”, „există autori care afirmă”, „se știe că”. În lucrările științifice trebuie **precizat cu acuratețe cine este autorul** conceptelor, argumentelor, formulărilor invocate⁴.

³ Earl Babbie, *Survey Research Methods*, Belmont: Wadsworth Publishing Company, 1990, pp. 480-482.

⁴ Cristian Pârvolescu (coord.), *Ghid de elaborare a lucrării de licență și a disertației*, București: Școala Națională de Studii Politice și Administrative, pp. 58-59, <http://www.politice.ro/img/ghidelaborare.pdf>, accesat în 12.06.2015.

„Nu am vrut să plagiez!”

Succint formulat, „plagiatul nu se măsoară prin intenția autorului, ci prin creația rezultată, mai precis, prin mesajul transmis publicului”⁵. Altfel spus, problema nu este aceea dacă studentul-doctorand a comis plagiatul intenționat sau nu, ci efectul produs de acesta, anume încălcarea dreptului de autor și/sau nerespectarea normelor de etică ale instituției care girează lucrarea respectivă.

Pentru ca studenții-doctoranzi să evite plagiatul, ceea ce trebuie să facă atunci **când conspectează** este:

- să citeze sursa, folosind ghilimelele și precizând referința;
- să nu copieze direct paragrafe, dacă nu se dorește folosirea lor ca atare, sub formă de citat;
- să rezume informațiile regăsite în text indicând referința;
- pentru fotografii, grafice și interpretări ale acestora trebuie să se ceară permisiunea de a le folosi. De regulă, autorii oferă fără rezerve acest lucru⁶.

În acest mod, sursele informațiilor folosite în redactarea lucrării sunt transparente, se respectă normele deontologice în vigoare și se asigură onestitatea intelectuală a autorului.

Exemplu

Text original:

„Security is taken to be about the pursuit of freedom from threat and the ability of states and societies to maintain their independent identity and their functional integrity against forces of change, which they see as hostile. The bottom line of security is survival, but it also reasonably includes a substantial range of concerns about the conditions of existence. Quite where this range of concerns ceases to merit the urgency of the “security” label (which identifies threats as significant enough to warrant emergency action and exceptional measures including

⁵ Ibidem, p. 59.

⁶ Alpár Szász, *Plagiatul: Forme și tehnici de evitare*, Facultatea de Științe Politice, Administrative și ale Comunicării, Universitatea Babeș-Bolyai, p. 2, <http://www.apubb.ro/wp-content/uploads/2011/03/ReguliPlagiat.pdf> accesat în 12.06.2015.

the use of force) and becomes part of everyday uncertainties of life is one of the difficulties of the concept”⁷.

Redare sub formă de citat:

„Se consideră că securitatea privește punerea la adăpost față de amenințări și capacitatea statelor și societăților de a-și menține identitatea independentă și integritatea lor funcțională, împotriva forțelor de schimbare, pe care le văd ca ostile. Aspectul esențial al securității este supraviețuirea, dar include, de asemenea, în mod rezonabil, o gamă substantială de preocupări cu privire la condițiile de existență. Unde anume această gamă de preocupări încetează să merite imperativul etichetei „securitate” (care identifică amenințările ca fiind suficient de semnificative pentru a justifica acțiuni de urgență și măsuri excepționale, inclusiv folosirea forței) și devine parte a incertitudinilor de zi cu zi reprezintă una dintre dificultățile conceptului”

Variantă concisă de redare:

Berry Buzan afirmă că securitatea vizează preocuparea statelor de a se proteja față de amenințări și de a-și menține identitatea, independența și integritatea funcțională. Nucleul conceptului de „securitate” este constituit de asigurarea supraviețuirii, dar demarcarea precisă a conceptului, dincolo de condițiile de asigurare a supraviețuirii, este problematică.

Alternativă de redare:

Berry Buzan consideră că securitatea presupune atât capacitatea statelor de a se proteja de amenințări, cât și de a-și menține identitatea și integritatea funcțională. Securitatea vizează, în principal, supraviețuirea, dar cuprinde și o serie de preocupări rezonabile legate de condițiile de existență ale statelor. Conform lui Buzan, este dificil de stabilit demarcația dintre preocupările care atentează la condițiile rezonabile de existență ale statelor și justifică intervenții excepționale din partea acestora și preocupările care țin de incertitudinile cotidiene cu care statele se confruntă.

Date despre subiecții intervievați

⁷ Barry Buzan, „New Patterns of Global Security in the Twenty-first Century” în *International Affairs*, Vol. 67 No. 3, 1991, pp. 432-433.

Două chestiuni se cuvin a fi menționate în ce privește prezentarea datelor despre subiecții intervievați:

- a) În cadrul tezei de doctorat, cercetările, mai ales cele calitative, presupun în mod uzual interviuarea unor subiecți. Astfel, este important de precizat că aceștia **au dreptul la anonim**, mai ales când este vorba de lucrări care se bazează pe un număr mic de subiecți ușor identificabili.
- b) În situația în care este necesară definirea unui atribut care ajută la identificarea unui subiect (de exemplu, un manager de companie, un șef de filială de partid, un înalt funcționar de stat) este obligatoriu ca **persoana în cauză să fie informată** despre faptul că în lucrarea elaborată vor fi utilizate date sau citate furnizate de acesta⁸.

Concluzie

Teza de doctorat este relevantă atât din punct de vedere științific cât și profesional, deoarece reprezintă o carte de vizită pentru cariera viitoare a absolvenților. Din acest motiv, respectarea normelor deontologice este crucială.

Rezumat

1. Facultatea de Studii Europene sancționează plagiatul prin respingerea lucrării.
2. Asigurați-vă că lucrarea nu se încadrează în vreunul dintre tipurile de plagiat:
 - (a) *Copiere integrală*
 - (b) *Copiere parțială*
 - (c) *Copiere prin parafrizare*
 - (d) *Auto-plagiere*
3. Evitați plagiatul prin cei patru pași menționați:

originalitate

credibilitate

citare

⁸ Cristian Părvulescu (coord.), *op. cit.*, p. 60.

acuratețe.

4. Formula „nu am vrut să copiez” nu disculpă.
5. Respectați dreptul la informare și la anonimat al subiecților intervievați.

ANEXE

Stilul Harvard de citare

1. Unul, doi sau trei autori, o carte:

- *în text:*
- (van Dalen 2008: 57)
- (Chiswell și Hodges 2007: 34)
- *la bibliografie:*
- van Dalen, Dirk. 2008. *Logic and Structure*. New York: Springer Verlag;
- Chiswell, Ian și Wilfrid Hodges. 2007. *Mathematical Logic*. Oxford: Oxford University Press.

2. Patru sau mai mulți autori, o carte – citați în text doar primul autor urmat de *et al.* iar la bibliografie citați toți autorii

- *în text:*
- (Laumann *et al.* 1994: 55)
- *la bibliografie:*
- Laumann, Edward O.; John H. Gagnon, Robert T. Michael și Stuart Michaels. 1994. *The social organization of sexuality: Sexual practices in the United States*. Chicago: University of Chicago Press.

3. Capitol sau articol dintr-o carte:

- *în text:*
- (Steiner 2005: 634)
- (Bueno 2009: 81)
- *la bibliografie:*

- Steiner, Mark. 2005. „Mathematics – Application and Applicability”. În Stewart Shapiro, editor, *The Oxford Handbook of Philosophy of Mathematics and Logic*, 624-649. Oxford: Oxford University Press.
- Bueno, Otávio. 2009. „Mathematical Fictionalism”. În Otávio Bueno, Øystein Linnebo, editori, *New Waves in Philosophy of Mathematics*, 53-83. Hampshire: Palgrave Macmillan.

4. Prefață, cuvânt înainte, introducere, studiu introductiv sau altă parte similară dintr-o carte

- *în text:*
- (Flonta 2008: 41)
- *la bibliografie:*
- Flonta, Mircea. 2008. Studiu introductiv la *Structura revoluțiilor științifice* de Thomas Kuhn, 5-49. București: Humanitas.

5. Cărți în format electronic (în cazul în care cartea este disponibilă în mai multe formate, citați formatul consultat. Pentru cărțile online specificați URL-ul; includeți data accesării doar în cazul în care citatul o cere – de ex. articol într-un cotidian. În cazul în care cartea nu are un sistem de numerotare a paginilor, includeți un titlu de secțiune sau de capitol)

- *în text:*
- (Boolos, Burgess și Jeffrey 2002: 127)
- *la bibliografie:*
- Boolos, George S.; John P. Burgess și Richard C. Jeffrey. 2002. *Computability and Logic*. Ediția a IV-a. New York: Cambridge University Press. Ediția Kindle.

6. Articol într-o revistă sau jurnal științific

- *în text:*
- (Smith, Button 2012: 115)

- *la bibliografie:*
- Smith, Peter și Tim Button. 2012. „The Philosophical Significance of Tennenbaum's Theorem”. *Philosophia Mathematica* 20 (1): 114-121.

7. Articol într-o revistă sau jurnal științific în format electronic

- *în text:*
- (Goodstein 1944)
- *la bibliografie:*
- Goodstein, Reuben. 1944. „On the restricted ordinal theorem”. *Journal of Symbolic Logic*, 9 (2): 33-41. doi: [10.2307/2268019](https://doi.org/10.2307/2268019). Accesat în 14.02.2015.

8. Articol într-un ziar sau cotidian (dacă articolul nu are autor, treceti doar titlul. Includeți un URL și data accesării, dacă citarea se face în urma consultării unei ediții electronice)

- *în text:*
- (MacFarquhar 2015)
- *la bibliografie:*
- MacFarquhar, Neil. 2015. „For Russia, Links Between Caucasus and ISIS Provoke Anxiety”. *New York Times*, 20 noiembrie.
<http://www.nytimes.com/pages/world/europe/index.html?action=click&contentCollection=Europe&module=Kicker®ion=Header&pgtype=article>. Accesat în 20.11.2015.

9. Teză de doctorat lucrare de disertație, lucrare de licență

- *în text* (Ludușan 2012)
- *la bibliografie:* Ludușan, Adrian. *Teorii ale referinței*. Teză de doctorat. Universitatea Babeș-Bolyai.

UNIVERSITATEA “BABEȘ-BOLYAI” CLUJ-NAPOCA
FACULTATEA DE STUDII EUROPENE
ȘCOALA DOCTORALĂ „PARADIGMA EUROPEANĂ”

TEZĂ DE DOCTORAT

Coordonator științific:
Titlu Prenume Nume

Absolvent:
Prenume Nume

Anul

UNIVERSITATEA BABEȘ-BOLYAI CLUJ-NAPOCA
FACULTATEA DE STUDII EUROPENE
ȘCOALA DOCTORALĂ „PARADIGMA EUROPEANĂ”

Titlul tezei de doctorat

Coordonator științific:
Titlu Prenume Nume

Absolvent:
Prenume Nume

Localitatea
Anul

Declarație

Prin prezenta declar că teza de doctorat cu titlul *Titlul complet al lucrării* este scrisă de mine și nu a mai fost prezentată niciodată la o altă facultate sau instituție de învățământ superior din țară sau străinătate. De asemenea, declar că toate sursele utilizate, inclusive cele de pe Internet, sunt indicate în lucrare, cu respectarea regulilor de evitare a plagiatului:

- toate fragmentele de text reproduse exact, chiar și în traducere proprie din altă limbă, sunt scrise între ghilimele și dețin referința precisă a sursei;
- reformularea în cuvinte proprii a textelor scrise de către alți autori deține referința precisă;
- rezumarea ideilor altor autori deține referința precisă la textul original.

Cluj-Napoca, *data*

Absolvent *Prenume Nume*

(semnătura olograf)

Cuprins

Introducere	45
1.Nume Capitol	46
1.1.Nume Subcapitol	46
1.2.Nume Subcapitol	47
2.Nume Capitol	48
2.1.Nume Subcapitol	48
2.2.Nume Subcapitol	48
2.3.Nume Subcapitol	48
3.Nume Capitol	49
3.1.Nume Subcapitol	49
3.2.Nume Subcapitol	49
3.3.Nume Subcapitol	49
4. etc.	50
Concluzii	51
Bibliografie	52

Introducere

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

1. Nume Capitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl. Pellentesque facilisis diam commodo quam sagittis lacinia. Vivamus a dui at dolor cursus egestas. Mauris efficitur arcu non odio malesuada tincidunt eu eget lorem. Duis eu ante lorem.

1.1. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

„Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl. Pellentesque facilisis diam commodo quam sagittis lacinia. Vivamus a dui at dolor cursus egestas. Mauris efficitur arcu non odio malesuada tincidunt eu eget lorem. Duis eu ante lorem.”⁹

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit.

⁹ Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl. Pellentesque facilisis diam commodo quam sagittis lacinia. Vivamus a dui at dolor cursus egestas. Mauris efficitur arcu non odio malesuada tincidunt eu eget lorem. Duis eu ante lorem.

1.2. Nume Subcapitol

Etc.

2. Nume Capitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

2.1. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

2.2. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

2.3. Nume Subcapitol

Etc.

3. Nume Capitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

3.1. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

3.2. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

3.3. Nume Subcapitol

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

4. etc.

Concluzii

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nunc quis faucibus urna. Etiam vestibulum sapien nisl, sit amet luctus tellus luctus sed. Morbi at condimentum mi, a convallis sapien. Cras blandit arcu in volutpat facilisis. Pellentesque feugiat congue nisi, vitae tempus orci placerat ac. Praesent erat ante, consectetur eu viverra sed, tincidunt vitae elit. Integer id rutrum metus, ac pellentesque libero. Curabitur blandit, tellus varius porttitor sagittis, lectus purus dapibus ex, vel facilisis nunc augue vitae nisl.

Bibliografie

SOCIAL SCIENCE RESEARCH: PRINCIPLES, METHODS, AND PRACTICES

Anol Bhattacharjee, Ph.D.

University of South Florida
Tampa, Florida, USA
abhattach@usf.edu

Second Edition
Copyright © 2012 by Anol Bhattacharjee

Published under the Creative Commons Attribution 3.0 License



Chapter 5

Research Design

Research design is a comprehensive plan for data collection in an empirical research project. It is a "blueprint" for empirical research aimed at answering specific research questions or testing specific hypotheses, and must specify at least three processes: (1) the data collection process, (2) the instrument development process, and (3) the sampling process. The instrument development and sampling processes are described in next two chapters, and the data collection process (which is often loosely called "research design") is introduced in this chapter and is described in further detail in Chapters 9-12.

Broadly speaking, data collection methods can be broadly grouped into two categories: positivist and interpretive. **Positivist methods**, such as laboratory experiments and survey research, are aimed at theory (or hypotheses) testing, while interpretive methods, such as action research and ethnography, are aimed at theory building. Positivist methods employ a deductive approach to research, starting with a theory and testing theoretical postulates using empirical data. In contrast, **interpretive methods** employ an inductive approach that starts with data and tries to derive a theory about the phenomenon of interest from the observed data. Often times, these methods are incorrectly equated with quantitative and qualitative research. Quantitative and qualitative methods refers to the type of data being collected (quantitative data involve numeric scores, metrics, and so on, while qualitative data includes interviews, observations, and so forth) and analyzed (i.e., using quantitative techniques such as regression or qualitative techniques such as coding). Positivist research uses predominantly quantitative data, but can also use qualitative data. Interpretive research relies heavily on qualitative data, but can sometimes benefit from including quantitative data as well. Sometimes, joint use of qualitative and quantitative data may help generate unique insight into a complex social phenomenon that are not available from either types of data alone, and hence, **mixed-mode designs** that combine qualitative and quantitative data are often highly desirable.

Key Attributes of a Research Design

The quality of research designs can be defined in terms of four key design attributes: internal validity, external validity, construct validity, and statistical conclusion validity.

Internal validity, also called causality, examines whether the observed change in a dependent variable is indeed caused by a corresponding change in hypothesized independent variable, and not by variables extraneous to the research context. Causality requires three conditions: (1) covariation of cause and effect (i.e., if cause happens, then effect also happens; and if cause does not happen, effect does not happen), (2) temporal precedence: cause must

precede effect in time, (3) no plausible alternative explanation (or spurious correlation). Certain research designs, such as laboratory experiments, are strong in internal validity by virtue of their ability to manipulate the independent variable (cause) via a treatment and observe the effect (dependent variable) of that treatment after a certain point in time, while controlling for the effects of extraneous variables. Other designs, such as field surveys, are poor in internal validity because of their inability to manipulate the independent variable (cause), and because cause and effect are measured at the same point in time which defeats temporal precedence making it equally likely that the expected effect might have influenced the expected cause rather than the reverse. Although higher in internal validity compared to other methods, laboratory experiments are, by no means, immune to threats of internal validity, and are susceptible to history, testing, instrumentation, regression, and other threats that are discussed later in the chapter on experimental designs. Nonetheless, different research designs vary considerably in their respective level of internal validity.

External validity or generalizability refers to whether the observed associations can be generalized from the sample to the population (population validity), or to other people, organizations, contexts, or time (ecological validity). For instance, can results drawn from a sample of financial firms in the United States be generalized to the population of financial firms (population validity) or to other firms within the United States (ecological validity)? Survey research, where data is sourced from a wide variety of individuals, firms, or other units of analysis, tends to have broader generalizability than laboratory experiments where artificially contrived treatments and strong control over extraneous variables render the findings less generalizable to real-life settings where treatments and extraneous variables cannot be controlled. The variation in internal and external validity for a wide range of research designs are shown in Figure 5.1.

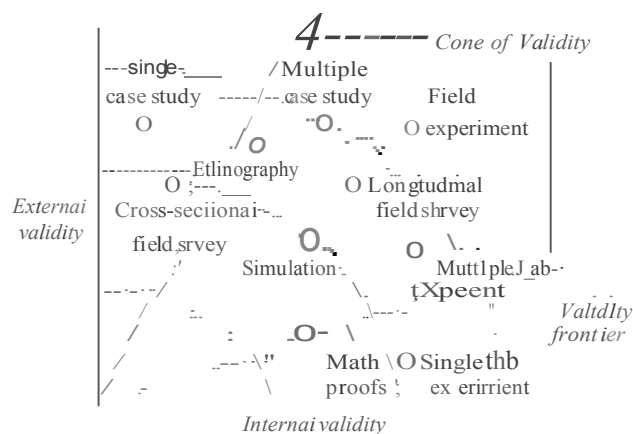


Figure 5.1. Internal and external validity

Some researchers claim that there is a tradeoff between internal and external validity: higher external validity can come only at the cost of internal validity and vice-versa. But this is not always the case. Research designs such as field experiments, longitudinal field surveys, and multiple case studies have higher degrees of both internal and external validities. Personally, I prefer research designs that have reasonable degrees of both internal and external validities, i.e., those that fall within the cone of validity shown in Figure 5.1. But this should not suggest that designs outside this cone are any less useful or valuable. Researchers' choice of designs is

ultimately a matter of their personal preference and competence, and the level of internal and external validity they desire.

Construct validity examines how well a given measurement scale is measuring the theoretical construct that it is expected to measure. Many constructs used in social science research such as empathy, resistance to change, and organizational learning are difficult to define, much less measure. For instance, construct validity must assure that a measure of empathy is indeed measuring empathy and not compassion, which may be difficult since these constructs are somewhat similar in meaning. Construct validity is assessed in positivist research based on correlational or factor analysis of pilot test data, as described in the next chapter.

Statistical conclusion validity examines the extent to which conclusions derived using a statistical procedure is valid. For example, it examines whether the right statistical method was used for hypotheses testing, whether the variables used meet the assumptions of that statistical test (such as sample size or distributional requirements), and so forth. Because interpretive research designs do not employ statistical test, statistical conclusion validity is not applicable for such analysis. The different kinds of validity and where they exist at the theoretical/empirical levels are illustrated in Figure 5.2.

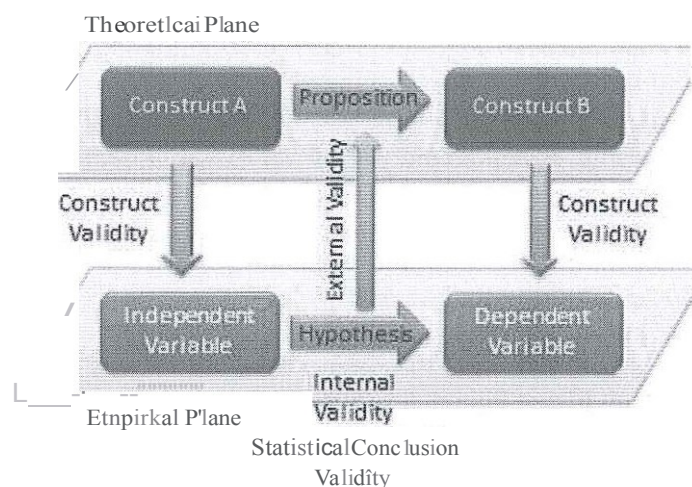


Figure 5.2. Different Types of Validity in Scientific Research

Improving Internal and External Validity

The best research designs are those that can assure high levels of internal and external validity. Such designs would guard against spurious correlations, inspire greater faith in the hypotheses testing, and ensure that the results drawn from a small sample are generalizable to the population at large. Controls are required to assure internal validity (causality) of research designs, and can be accomplished in four ways: (1) manipulation, (2) elimination, (3) inclusion, and (4) statistical control, and (5) randomization.

In **manipulation**, the researcher manipulates the independent variables in one or more levels (called "treatments"), and compares the effects of the treatments against a control group where subjects do not receive the treatment. Treatments may include a new drug or different

dosage of drug (for treating a medical condition), a teaching style (for students), and so forth. This type of control is achieved in experimental or quasi-experimental designs but not in non-experimental designs such as surveys. Note that if subjects cannot distinguish adequately between different levels of treatment manipulations, their responses across treatments may not be different, and manipulation would fail.

The **elimination** technique relies on eliminating extraneous variables by holding them constant across treatments, such as by restricting the study to a single gender or a single socioeconomic status. In the **inclusion** technique, the role of extraneous variables is considered by including them in the research design and separately estimating their effects on the dependent variable, such as via factorial designs where one factor is gender (male versus female). Such technique allows for greater generalizability but also requires substantially larger samples. In **statistical control**, extraneous variables are measured and used as covariates during the statistical testing process.

Finally, the **randomization** technique is aimed at canceling out the effects of extraneous variables through a process of random sampling, if it can be assured that these effects are of a random (non-systematic) nature. Two types of randomization are: (1) **random selection**, where a sample is selected randomly from a population, and (2) **random assignment**, where subjects selected in a non-random manner are randomly assigned to treatment groups.

Randomization also assures external validity, allowing inferences drawn from the sample to be generalized to the population from which the sample is drawn. Note that random assignment is mandatory when random selection is not possible because of resource or access constraints. However, generalizability across populations is harder to ascertain since populations may differ on multiple dimensions and you can only control for few of those dimensions.

Popular Research Designs

As noted earlier, research designs can be classified into two categories – positivist and interpretive – depending on their goal in scientific research. Positivist designs are meant for theory testing, while interpretive designs are meant for theory building. Positivist designs seek generalized patterns based on an objective view of reality, while interpretive designs seek subjective interpretations of social phenomena from the perspectives of the subjects involved. Some popular examples of positivist designs include laboratory experiments, field experiments, field surveys, secondary data analysis, and case research while examples of interpretive designs include case research, phenomenology, and ethnography. Note that case research can be used for theory building or theory testing, though not at the same time. Not all techniques are suited for all kinds of scientific research. Some techniques such as focus groups are best suited for exploratory research, others such as ethnography are best for descriptive research, and still others such as laboratory experiments are ideal for explanatory research. Following are brief descriptions of some of these designs. Additional details are provided in Chapters 9-12.

Experimental studies are those that are intended to test cause-effect relationships (hypotheses) in a tightly controlled setting by separating the cause from the effect in time, administering the cause to one group of subjects (the "treatment group") but not to another group ("control group"), and observing how the mean effects vary between subjects in these two groups. For instance, if we design a laboratory experiment to test the efficacy of a new drug in treating a certain ailment, we can get a random sample of people afflicted with that ailment,

randomly assign them to one of two groups (treatment and control groups), administer the drug to subjects in the treatment group, but only give a placebo (e.g., a sugar pill with no medicinal value). More complex designs may include multiple treatment groups, such as low versus high dosage of the drug, multiple treatments, such as combining drug administration with dietary interventions. In a **true experimental design**, subjects must be randomly assigned between each group. If random assignment is not followed, then the design becomes **quasi-experimental**. Experiments can be conducted in an artificial or laboratory setting such as at a university (laboratory experiments) or in field settings such as in an organization where the phenomenon of interest is actually occurring (field experiments). Laboratory experiments allow the researcher to isolate the variables of interest and control for extraneous variables, which may not be possible in field experiments. Hence, inferences drawn from laboratory experiments tend to be stronger in internal validity, but those from field experiments tend to be stronger in external validity. Experimental data is analyzed using quantitative statistical techniques. The primary strength of the experimental design is its strong internal validity due to its ability to isolate, control, and intensively examine a small number of variables, while its primary weakness is limited external generalizability since real life is often more complex (i.e., involve more extraneous variables) than contrived lab settings. Furthermore, if the research does not identify ex ante relevant extraneous variables and control for such variables, such lack of controls may hurt internal validity and may lead to spurious correlations.

Field surveys are non-experimental designs that do not control for or manipulate independent variables or treatments, but measure these variables and test their effects using statistical methods. Field surveys capture snapshots of practices, beliefs, or situations from a random sample of subjects in field settings through a survey questionnaire or less frequently, through a structured interview. In **cross-sectional field surveys**, independent and dependent variables are measured at the same point in time (e.g., using a single questionnaire), while in **longitudinal field surveys**, dependent variables are measured at a later point in time than the independent variables. The strengths of field surveys are their external validity (since data is collected in field settings), their ability to capture and control for a large number of variables, and their ability to study a problem from multiple perspectives or using multiple theories. However, because of their non-temporal nature, internal validity (cause-effect relationships) are difficult to infer, and surveys may be subject to respondent biases (e.g., subjects may provide a "socially desirable" response rather than their true response) which further hurts internal validity.

Secondary data analysis is an analysis of data that has previously been collected and tabulated by other sources. Such data may include data from government agencies such as employment statistics from the U.S. Bureau of Labor Services or development statistics by country from the United Nations Development Program, data collected by other researchers (often used in meta-analytic studies), or publicly available third-party data, such as financial data from stock markets or real-time auction data from eBay. This is in contrast to most other research designs where collecting primary data for research is part of the researcher's job. Secondary data analysis may be an effective means of research where primary data collection is too costly or infeasible, and secondary data is available at a level of analysis suitable for answering the researcher's questions. The limitations of this design are that the data might not have been collected in a systematic or scientific manner and hence unsuitable for scientific research, since the data was collected for a presumably different purpose, they may not adequately address the research questions of interest to the researcher, and internal validity is problematic if the temporal precedence between cause and effect is unclear.

Case research is an in-depth investigation of a problem in one or more real-life settings (case sites) over an extended period of time. Data may be collected using a combination of interviews, personal observations, and internal or external documents. Case studies can be positivist in nature (for hypotheses testing) or interpretive (for theory building). The strength of this research method is its ability to discover a wide variety of social, cultural, and political factors potentially related to the phenomenon of interest that may not be known in advance. Analysis tends to be qualitative in nature, but heavily contextualized and nuanced. However, interpretation of findings may depend on the observational and integrative ability of the researcher, lack of control may make it difficult to establish causality, and findings from a single case site may not be readily generalized to other case sites. Generalizability can be improved by replicating and comparing the analysis in other case sites in a **multiple case design**.

Focus group research is a type of research that involves bringing in a small group of subjects (typically 6 to 10 people) at one location, and having them discuss a phenomenon of interest for a period of 1.5 to 2 hours. The discussion is moderated and led by a trained facilitator, who sets the agenda and poses an initial set of questions for participants, makes sure that ideas and experiences of all participants are represented, and attempts to build a holistic understanding of the problem situation based on participants' comments and experiences. Internal validity cannot be established due to lack of controls and the findings may not be generalized to other settings because of small sample size. Hence, focus groups are not generally used for explanatory or descriptive research, but are more suited for exploratory research.

Action research assumes that complex social phenomena are best understood by introducing interventions or "actions" into those phenomena and observing the effects of those actions. In this method, the researcher is usually a consultant or an organizational member embedded within a social context such as an organization, who initiates an action such as new organizational procedures or new technologies, in response to a real problem such as declining profitability or operational bottlenecks. The researcher's choice of actions must be based on theory, which should explain why and how such actions may cause the desired change. The researcher then observes the results of that action, modifying it as necessary, while simultaneously learning from the action and generating theoretical insights about the target problem and interventions. The initial theory is validated by the extent to which the chosen action successfully solves the target problem. Simultaneous problem solving and insight generation is the central feature that distinguishes action research from all other research methods, and hence, action research is an excellent method for bridging research and practice. This method is also suited for studying unique social problems that cannot be replicated outside that context, but it is also subject to researcher bias and subjectivity, and the generalizability of findings is often restricted to the context where the study was conducted.

Ethnography is an interpretive research design inspired by anthropology that emphasizes that research phenomenon must be studied within the context of its culture. The researcher is deeply immersed in a certain culture over an extended period of time (8 months to 2 years), and during that period, engages, observes, and records the daily life of the studied culture, and theorizes about the evolution and behaviors in that culture. Data is collected primarily via observational techniques, formal and informal interaction with participants in that culture, and personal field notes, while data analysis involves "sense-making". The researcher must narrate her experience in great detail so that readers may experience that same culture without necessarily being there. The advantages of this approach are its sensitiveness to the context, the rich and nuanced understanding it generates, and minimal

respondent bias. However, this is also an extremely time and resource-intensive approach, and findings are specific to a given culture and less generalizable to other cultures.

Selecting Research Designs

Given the above multitude of research designs, which design should researchers choose for their research? Generally speaking, researchers tend to select those research designs that they are most comfortable with and feel most competent to handle, but ideally, the choice should depend on the nature of the research phenomenon being studied. In the preliminary phases of research, when the research problem is unclear and the researcher wants to scope out the nature and extent of a certain research problem, a focus group (for individual unit of analysis) or a case study (for organizational unit of analysis) is an ideal strategy for exploratory research. As one delves further into the research domain, but finds that there are no good theories to explain the phenomenon of interest and wants to build a theory to fill in the unmet gap in that area, interpretive designs such as case research or ethnography may be useful designs. If competing theories exist and the researcher wishes to test these different theories or integrate them into a larger theory, positivist designs such as experimental design, survey research, or secondary data analysis are more appropriate.

Regardless of the specific research design chosen, the researcher should strive to collect quantitative and qualitative data using a combination of techniques such as questionnaires, interviews, observations, documents, or secondary data. For instance, even in a highly structured survey questionnaire, intended to collect quantitative data, the researcher may leave some room for a few open-ended questions to collect qualitative data that may generate unexpected insights not otherwise available from structured quantitative data alone. Likewise, while case research employs mostly face-to-face interviews to collect mostly qualitative data, the potential and value of collecting quantitative data should not be ignored. As an example, in a study of organizational decision making processes, the case interviewer can record numeric quantities such as how many months it took to make certain organizational decisions, how many people were involved in that decision process, and how many decision alternatives were considered, which can provide valuable insights not otherwise available from interviewees' narrative responses. Irrespective of the specific research design employed, the goal of the researcher should be to collect as much and as diverse data as possible that can help generate the best possible insights about the phenomenon of interest.

SOCIAL SCIENCE RESEARCH: PRINCIPLES, METHODS, AND PRACTICES

Anol Bhattacharjee, Ph.D.

University of South Florida
Tampa, Florida, USA
abhattach@usf.edu

Second Edition
Copyright © 2012 by Anol Bhattacharjee

Published under the Creative Commons Attribution 3.0 License



Chapter 3

The Research Process

In Chapter 1, we saw that scientific research is the process of acquiring scientific knowledge using the scientific method. But how is such research conducted? This chapter delves into the process of scientific research, and the assumptions and outcomes of the research process.

Paradigms of Social Research

Our design and conduct of research is shaped by our mental models or frames of references that we use to organize our reasoning and observations. These mental models or frames (belief systems) are called **paradigms**. The word "paradigm" was popularized by Thomas Kuhn (1962) in his book *The Structure of Scientific Revolutions*, where he examined the history of the natural sciences to identify patterns of activities that shape the progress of science. Similar ideas are applicable to social sciences as well, where a social reality can be viewed by different people in different ways, which may constrain their thinking and reasoning about the observed phenomenon. For instance, conservatives and liberals tend to have very different perceptions of the role of government in people's lives, and hence, have different opinions on how to solve social problems. Conservatives may believe that lowering taxes is the best way to stimulate a stagnant economy because it increases people's disposable income and spending, which in turn expands business output and employment. In contrast, liberals may believe that governments should invest more directly in job creation programs such as public works and infrastructure projects, which will increase employment and people's ability to consume and drive the economy. Likewise, Western societies place greater emphasis on individual rights, such as one's right to privacy, right of free speech, and right to bear arms. In contrast, Asian societies tend to balance the rights of individuals against the rights of families, organizations, and the government, and therefore tend to be more communal and less individualistic in their policies. Such differences in perspective often lead Westerners to criticize Asian governments for being autocratic, while Asians criticize Western societies for being greedy, having high crime rates, and creating a "cult of the individual." Our personal paradigms are like "colored glasses" that govern how we view the world and how we structure our thoughts about what we see in the world.

Paradigms are often hard to recognize, because they are implicit, assumed, and taken for granted. However, recognizing these paradigms is key to making sense of and reconciling differences in people's perceptions of the same social phenomenon. For instance, why do liberals believe that the best way to improve secondary education is to hire more teachers, but conservatives believe that privatizing education (using such means as school vouchers) are

more effective in achieving the same goal? Because conservatives place more faith in competitive markets (i.e., in free competition between schools competing for education dollars), while liberals believe more in labor (i.e., in having more teachers and schools). Likewise, in social science research, if one were to understand why a certain technology was successfully implemented in one organization but failed miserably in another, a researcher looking at the world through a "rational lens" will look for rational explanations of the problem such as inadequate technology or poor fit between technology and the task context where it is being utilized, while another researcher looking at the same problem through a "social lens" may seek out social deficiencies such as inadequate user training or lack of management support, while those seeing it through a "political lens" will look for instances of organizational politics that may subvert the technology implementation process. Hence, subconscious paradigms often constrain the concepts that researchers attempt to measure, their observations, and their subsequent interpretations of a phenomenon. However, given the complex nature of social phenomenon, it is possible that all of the above paradigms are partially correct, and that a fuller understanding of the problem may require an understanding and application of multiple paradigms.

Two popular paradigms today among social science researchers are positivism and post-positivism. **Positivism**, based on the works of French philosopher Auguste Comte (1798- 1857), was the dominant scientific paradigm until the mid-20th century. It holds that science or knowledge creation should be restricted to what can be observed and measured. Positivism tends to rely exclusively on theories that can be directly tested. Though positivism was originally an attempt to separate scientific inquiry from religion (where the precepts could not be objectively observed), positivism led to *empiricism* or a blind faith in observed data and a rejection of any attempt to extend or reason beyond observable facts. Since human thoughts and emotions could not be directly measured, there were not considered to be legitimate topics for scientific research. Frustrations with the strictly empirical nature of positivist philosophy led to the development of **post-positivism** (or postmodernism) during the mid-late 20th century. Post-positivism argues that one can make reasonable inferences about a phenomenon by combining empirical observations with logical reasoning. Post-positivists view science as not certain but probabilistic (i.e., based on many contingencies), and often seek to explore these contingencies to understand social reality better. The post-positivist camp has further fragmented into *subjectivists*, who view the world as a subjective construction of our subjective minds rather than as an objective reality, and *critical realists*, who believe that there is an external reality that is independent of a person's thinking but we can never know such reality with any degree of certainty.

Burrell and Morgan (1979), in their seminal book *Sociological Paradigms and Organizational Analysis*, suggested that the way social science researchers view and study social phenomena is shaped by two fundamental sets of philosophical assumptions: ontology and epistemology. **Ontology** refers to our assumptions about how we see the world, e.g., does the world consist mostly of social order or constant change. **Epistemology** refers to our assumptions about the best way to study the world, e.g., should we use an objective or subjective approach to study social reality. Using these two sets of assumptions, we can categorize social science research as belonging to one of four categories (see Figure 3.1).

If researchers view the world as consisting mostly of social order (ontology) and hence seek to study patterns of ordered events or behaviors, and believe that the best way to study such a world is using objective approach (epistemology) that is independent of the person conducting the observation or interpretation, such as by using standardized data collection

tools like surveys, then they are adopting a paradigm of **functionalism**. However, if they believe that the best way to study social order is through the subjective interpretation of participants involved, such as by interviewing different participants and reconciling differences among their responses using their own subjective perspectives, then they are employing an **interpretivism** paradigm. If researchers believe that the world consists of radical change and seek to understand or enact change using an objectivist approach, then they are employing a **radical structuralism** paradigm. If they wish to understand social change using the subjective perspectives of the participants involved, then they are following a **radical humanism** paradigm.

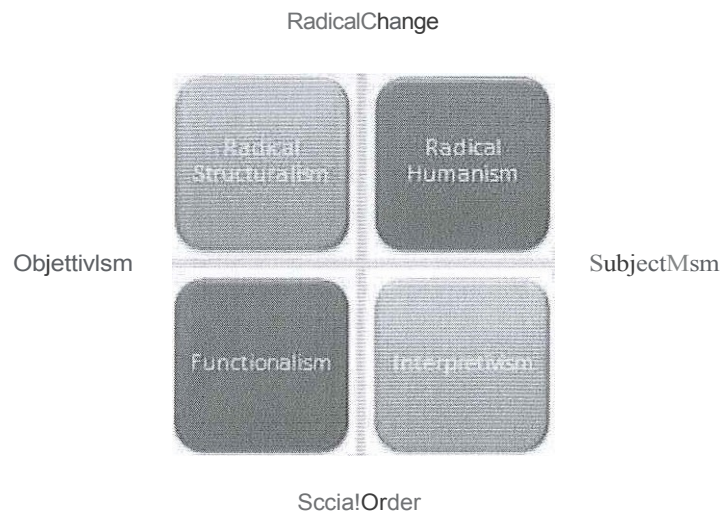


Figure 3.1 Four paradigms of social science research
(Source: Burrell and Morgan, 1979)

To date, the majority of social science research has emulated the natural sciences, and followed the functionalist paradigm. Functionalists believe that social order or patterns can be understood in terms of their functional components, and therefore attempt to break down a problem into small components and studying one or more components in detail using objectivist techniques such as surveys and experimental research. However, with the emergence of post-positivist thinking, a small but growing number of social science researchers are attempting to understand social order using subjectivist techniques such as interviews and ethnographic studies. Radical humanism and radical structuralism continues to represent a negligible proportion of social science research, because scientists are primarily concerned with understanding generalizable patterns of behavior, events, or phenomena, rather than idiosyncratic or changing events. Nevertheless, if you wish to study social change, such as why democratic movements are increasingly emerging in Middle Eastern countries, or why this movement was successful in Tunisia, took a longer path to success in Libya, and is still not successful in Syria, then perhaps radical humanism is the right approach for such a study. Social and organizational phenomena generally consists elements of both order and change. For instance, organizational success depends on formalized business processes, work procedures, and job responsibilities, while being simultaneously constrained by a constantly changing mix of competitors, competing products, suppliers, and customer base in the business environment. Hence, a holistic and more complete understanding of social phenomena such as why are some organizations more successful than others, require an appreciation and application of a multi-paradigmatic approach to research.

Overview of the Research Process

So how do our mental paradigms shape social science research? At its core, all scientific research is an iterative process of observation, rationalization, and validation. In the observation phase, we observe a natural or social phenomenon, event, or behavior that interests us. In the rationalization phase, we try to make sense of or the observed phenomenon, event, or behavior by logically connecting the different pieces of the puzzle that we observe, which in some cases, may lead to the construction of a theory. Finally, in the validation phase, we test our theories using a scientific method through a process of data collection and analysis, and in doing so, possibly modify or extend our initial theory. However, research designs vary based on whether the researcher starts at observation and attempts to rationalize the observations (inductive research), or whether the researcher starts at an ex ante rationalization or a theory and attempts to validate the theory (deductive research). Hence, the observation-rationalization-validation cycle is very similar to the induction-deduction cycle of research discussed in Chapter 1.

Most traditional research tends to be deductive and functionalistic in nature. Figure 3.2 provides a schematic view of such a research project. This figure depicts a series of activities to be performed in functionalist research, categorized into three phases: exploration, research design, and research execution. Note that this generalized design is not a roadmap or flowchart for all research. It applies only to functionalistic research, and it can and should be modified to fit the needs of a specific project.

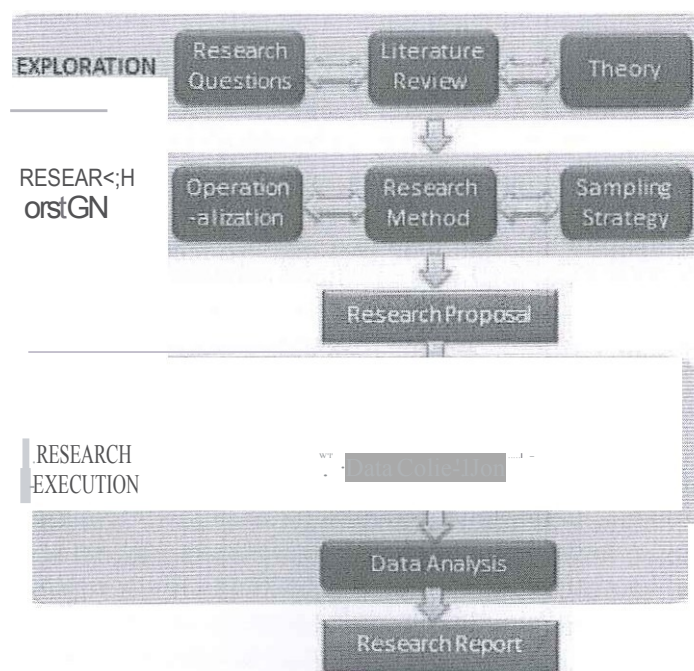


Figure 3.2. Functionalistic research process

The first phase of research is exploration. This phase includes exploring and selecting research questions for further investigation, examining the published literature in the area of inquiry to understand the current state of knowledge in that area, and identifying theories that may help answer the research questions of interest.

The first step in the exploration phase is identifying one or more **research questions** dealing with a specific behavior, event, or phenomena of interest. Research questions are specific questions about a behavior, event, or phenomena of interest that you wish to seek answers for in your research. Examples include what factors motivate consumers to purchase goods and services online without knowing the vendors of these goods or services, how can we make high school students more creative, and why do some people commit terrorist acts. Research questions can delve into issues of what, why, how, when, and so forth. More interesting research questions are those that appeal to a broader population (e.g., "how can firms innovate" is a more interesting research question than "how can Chinese firms innovate in the service-sector"), address real and complex problems (in contrast to hypothetical or "toy" problems), and where the answers are not obvious. Narrowly focused research questions (often with a binary yes/no answer) tend to be less useful and less interesting and less suited to capturing the subtle nuances of social phenomena. Uninteresting research questions generally lead to uninteresting and unpublishable research findings.

The next step is to conduct a **literature review** of the domain of interest. The purpose of a literature review is three-fold: (1) to survey the current state of knowledge in the area of inquiry, (2) to identify key authors, articles, theories, and findings in that area, and (3) to identify gaps in knowledge in that research area. Literature review is commonly done today using computerized keyword searches in online databases. Keywords can be combined using "and" and "or" operations to narrow down or expand the search results. Once a shortlist of relevant articles is generated from the keyword search, the researcher must then manually browse through each article, or at least its abstract section, to determine the suitability of that article for a detailed review. Literature reviews should be reasonably complete, and not restricted to a few journals, a few years, or a specific methodology. Reviewed articles may be summarized in the form of tables, and can be further structured using organizing frameworks such as a concept matrix. A well-conducted literature review should indicate whether the initial research questions have already been addressed in the literature (which would obviate the need to study them again), whether there are newer or more interesting research questions available, and whether the original research questions should be modified or changed in light of findings of the literature review. The review can also provide some intuitions or potential answers to the questions of interest and/or help identify theories that have previously been used to address similar questions.

Since functionalist (deductive) research involves theory-testing, the third step is to identify one or more **theories** that can help address the desired research questions. While the literature review may uncover a wide range of concepts or constructs potentially related to the phenomenon of interest, a theory will help identify which of these constructs is logically relevant to the target phenomenon and how. Forgoing theories may result in measuring a wide range of less relevant, marginally relevant, or irrelevant constructs, while also minimizing the chances of obtaining results that are meaningful and not by pure chance. In functionalist research, theories can be used as the logical basis for postulating hypotheses for empirical testing. Obviously, not all theories are well-suited for studying all social phenomena. Theories must be carefully selected based on their fit with the target problem and the extent to which their assumptions are consistent with that of the target problem. We will examine theories and the process of theorizing in detail in the next chapter.

The next phase in the research process is **research design**. This process is concerned with creating a blueprint of the activities to take in order to satisfactorily answer the research

questions identified in the exploration phase. This includes selecting a research method, operationalizing constructs of interest, and devising an appropriate sampling strategy.

Operationalization is the process of designing precise measures for abstract theoretical constructs. This is a major problem in social science research, given that many of the constructs, such as prejudice, alienation, and liberalism are hard to define, let alone measure accurately. Operationalization starts with specifying an "operational definition" (or "conceptualization") of the constructs of interest. Next, the researcher can search the literature to see if there are existing prevalidated measures matching their operational definition that can be used directly or modified to measure their constructs of interest. If such measures are not available or if existing measures are poor or reflect a different conceptualization than that intended by the researcher, new instruments may have to be designed for measuring those constructs. This means specifying exactly how exactly the desired construct will be measured (e.g., how many items, what items, and so forth). This can easily be a long and laborious process, with multiple rounds of pretests and modifications before the newly designed instrument can be accepted as "scientifically valid." We will discuss operationalization of constructs in a future chapter on measurement.

Simultaneously with operationalization, the researcher must also decide what **research method** they wish to employ for collecting data to address their research questions of interest. Such methods may include quantitative methods such as experiments or survey research or qualitative methods such as case research or action research, or possibly a combination of both. If an experiment is desired, then what is the experimental design? If survey, do you plan a mail survey, telephone survey, web survey, or a combination? For complex, uncertain, and multi-faceted social phenomena, multi-method approaches may be more suitable, which may help leverage the unique strengths of each research method and generate insights that may not be obtained using a single method.

Researchers must also carefully choose the target population from which they wish to collect data, and a **sampling** strategy to select a sample from that population. For instance, should they survey individuals or firms or workgroups within firms? What types of individuals or firms they wish to target? Sampling strategy is closely related to the unit of analysis in a research problem. While selecting a sample, reasonable care should be taken to avoid a biased sample (e.g., sample based on convenience) that may generate biased observations. Sampling is covered in depth in a later chapter.

At this stage, it is often a good idea to write a **research proposal** detailing all of the decisions made in the preceding stages of the research process and the rationale behind each decision. This multi-part proposal should address what research questions you wish to study and why, the prior state of knowledge in this area, theories you wish to employ along with hypotheses to be tested, how to measure constructs, what research method to be employed and why, and desired sampling strategy. Funding agencies typically require such a proposal in order to select the best proposals for funding. Even if funding is not sought for a research project, a proposal may serve as a useful vehicle for seeking feedback from other researchers and identifying potential problems with the research project (e.g., whether some important constructs were missing from the study) before starting data collection. This initial feedback is invaluable because it is often too late to correct critical problems after data is collected in a research study.

Having decided who to study (subjects), what to measure (concepts), and how to collect data (research method), the researcher is now ready to proceed to the **research execution** phase. This includes pilot testing the measurement instruments, data collection, and data analysis.

Pilot testing is an often overlooked but extremely important part of the research process. It helps detect potential problems in your research design and/or instrumentation (e.g., whether the questions asked are intelligible to the targeted sample), and to ensure that the measurement instruments used in the study are reliable and valid measures of the constructs of interest. The pilot sample is usually a small subset of the target population. After a successful pilot testing, the researcher may then proceed with **data collection** using the sampled population. The data collected may be quantitative or qualitative, depending on the research method employed.

Following data collection, the data is analyzed and interpreted for the purpose of drawing conclusions regarding the research questions of interest. Depending on the type of data collected (quantitative or qualitative), **data analysis** may be quantitative (e.g., employ statistical techniques such as regression or structural equation modeling) or qualitative (e.g., coding or content analysis).

The final phase of research involves preparing the final **research report** documenting the entire research process and its findings in the form of a research paper, dissertation, or monograph. This report should outline in detail all the choices made during the research process (e.g., theory used, constructs selected, measures used, research methods, sampling, etc.) and why, as well as the outcomes of each phase of the research process. The research process must be described in sufficient detail so as to allow other researchers to replicate your study, test the findings, or assess whether the inferences derived are scientifically acceptable. Of course, having a ready research proposal will greatly simplify and quicken the process of writing the finished report. Note that research is of no value unless the research process and outcomes are documented for future generations; such documentation is essential for the incremental progress of science.

Common Mistakes in Research

The research process is fraught with problems and pitfalls, and novice researchers often find, after investing substantial amounts of time and effort into a research project, that their research questions were not sufficiently answered, or that the findings were not interesting enough, or that the research was not of "acceptable" scientific quality. Such problems typically result in research papers being rejected by journals. Some of the more frequent mistakes are described below.

Insufficiently motivated research questions. Often times, we choose our "pet" problems that are interesting to us but not to the scientific community at large, i.e., it does not generate new knowledge or insight about the phenomenon being investigated. Because the research process involves a significant investment of time and effort on the researcher's part, the researcher must be certain (and be able to convince others) that the research questions they seek to answer in fact deal with real problems (and not hypothetical problems) that affect a substantial portion of a population and has not been adequately addressed in prior research.

Pursuing research fads. Another common mistake is pursuing "popular" topics with limited shelf life. A typical example is studying technologies or practices that are popular today. Because research takes several years to complete and publish, it is possible that popular interest in these fads may die down by the time the research is completed and submitted for publication. A better strategy may be to study "timeless" topics that have always persisted through the years.

Unresearchable problems. Some research problems may not be answered adequately based on observed evidence alone, or using currently accepted methods and procedures. Such problems are best avoided. However, some unresearchable, ambiguously defined problems may be modified or fine tuned into well-defined and useful researchable problems.

Favored research methods. Many researchers have a tendency to recast a research problem so that it is amenable to their favorite research method (e.g., survey research). This is an unfortunate trend. Research methods should be chosen to best fit a research problem, and not the other way around.

Blind data mining. Some researchers have the tendency to collect data first (using instruments that are already available), and then figure out what to do with it. Note that data collection is only one step in a long and elaborate process of planning, designing, and executing research. In fact, a series of other activities are needed in a research process prior to data collection. If researchers jump into data collection without such elaborate planning, the data collected will likely be irrelevant, imperfect, or useless, and their data collection efforts may be entirely wasted. An abundance of data cannot make up for deficits in research planning and design, and particularly, for the lack of interesting research questions.

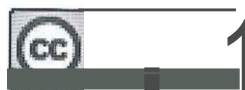
SOCIAL SCIENCE RESEARCH: PRINCIPLES, METHODS, AND PRACTICES

Anol Bhattacharjee, Ph.O.

University of South Florida
Tampa, Florida, USA
abha tt@usf .edu

Second Edition
Copyright © 2012 by Anol Bhattacharjee

Published under the Creative Commons Attribution 3.0 License



Chapter 2

Thinking Like a Researcher

Conducting good research requires first retraining your brain to think like a researcher. This requires visualizing the abstract from actual observations, mentally "connecting the dots" to identify hidden concepts and patterns, and synthesizing those patterns into generalizable laws and theories that apply to other contexts beyond the domain of the initial observations. Research involves constantly moving back and forth from an empirical plane where observations are conducted to a theoretical plane where these observations are abstracted into generalizable laws and theories. This is a skill that takes many years to develop, is not something that is taught in graduate or doctoral programs or acquired in industry training, and is by far the biggest deficit amongst Ph.D. students. Some of the mental abstractions needed to think like a researcher include unit of analysis, constructs, hypotheses, operationalization, theories, models, induction, deduction, and so forth, which we will examine in this chapter.

Unit of Analysis

One of the first decisions in any social science research is the unit of analysis of a scientific study. The **unit of analysis** refers to the person, collective, or object that is the target of the investigation. Typical units of analysis include individuals, groups, organizations, countries, technologies, objects, and such. For instance, if we are interested in studying people's shopping behavior, their learning outcomes, or their attitudes to new technologies, then the unit of analysis is the *individual*. If we want to study characteristics of street gangs or teamwork in organizations, then the unit of analysis is the *group*. If the goal of research is to understand how firms can improve profitability or make good executive decisions, then the unit of analysis is the *firm*. In this case, even though decisions are made by individuals in these firms, these individuals are presumed to represent their firm's decision rather than their personal decisions. If research is directed at understanding differences in national cultures, then the unit of analysis becomes a *country*. Even inanimate objects can serve as units of analysis. For instance, if a researcher is interested in understanding how to make web pages more attractive to its users, then the unit of analysis is a *web page* (and not users). If we wish to study how knowledge transfer occurs between two firms, then our unit of analysis becomes the *dyad* (the combination of firms that is sending and receiving knowledge).

Understanding the units of analysis can sometimes be fairly complex. For instance, if we wish to study why certain neighborhoods have high crime rates, then our unit of analysis becomes the *neighborhood*, and not crimes or criminals committing such crimes. This is because the object of our inquiry is the neighborhood and not criminals. However, if we wish to compare different types of crimes in different neighborhoods, such as homicide, robbery,

assault, and so forth, our unit of analysis becomes the *crime*. If we wish to study why criminals engage in illegal activities, then the unit of analysis becomes the *individual* (i.e., the criminal). Like, if we want to study why some innovations are more successful than others, then our unit of analysis is an *innovation*. However, if we wish to study how some organizations innovate more consistently than others, then the unit of analysis is the *organization*. Hence, two related research questions within the same research study may have two entirely different units of analysis.

Understanding the unit of analysis is important because it shapes what type of data you should collect for your study and who you collect it from. If your unit of analysis is a web page, you should be collecting data about web pages from actual web pages, and not surveying people about how they use web pages. If your unit of analysis is the organization, then you should be measuring organizational-level variables such as organizational size, revenues, hierarchy, or absorptive capacity. This data may come from a variety of sources such as financial records or surveys of Chief Executive Officers (CEO), who are presumed to be representing their organization (rather than themselves). Some variables such as CEO pay may seem like individual level variables, but in fact, it can also be an organizational level variable because each organization has only one CEO pay at any time. Sometimes, it is possible to collect data from a lower level of analysis and aggregate that data to a higher level of analysis. For instance, in order to study teamwork in organizations, you can survey individual team members in different organizational teams, and average their individual scores to create a composite team-level score for team-level variables like cohesion and conflict. We will examine the notion of "variables" in greater depth in the next section.

Concepts, Constructs, and Variables

We discussed in Chapter 1 that although research can be exploratory, descriptive, or explanatory, most scientific research tend to be of the explanatory type in that they search for potential explanations of observed natural or social phenomena. Explanations require development of **concepts** or generalizable properties or characteristics associated with objects, events, or people. While objects such as a person, a firm, or a car are not concepts, their specific characteristics or behavior such as a person's attitude toward immigrants, a firm's capacity for innovation, and a car's weight can be viewed as concepts.

Knowingly or unknowingly, we use different kinds of concepts in our everyday conversations. Some of these concepts have been developed over time through our shared language. Sometimes, we borrow concepts from other disciplines or languages to explain a phenomenon of interest. For instance, the idea of *gravitation* borrowed from physics can be used in business to describe why people tend to "gravitate" to their preferred shopping destinations. Likewise, the concept of *distance* can be used to explain the degree of social separation between two otherwise collocated individuals. Sometimes, we create our own concepts to describe a unique characteristic not described in prior research. For instance, *technostress* is a new concept referring to the mental stress one may face when asked to learn a new technology.

Concepts may also have progressive levels of abstraction. Some concepts such as a person's *weight* are precise and objective, while other concepts such as a person's *personality* may be more abstract and difficult to visualize. A **construct** is an abstract concept that is specifically chosen (or "created") to explain a given phenomenon. A construct may be a simple concept, such as a person's *weight*, or a combination of a set of related concepts such as a

person's *communication skill*, which may consist of several underlying concepts such as the person's *vocabulary*, *syntax*, and *spelling*. The former instance (weight) is a **unidimensional construct**, while the latter (communication skill) is a **multi-dimensional construct** (i.e., it consists of multiple underlying concepts). The distinction between constructs and concepts are clearer in multi-dimensional constructs, where the higher order abstraction is called a construct and the lower order abstractions are called concepts. However, this distinction tends to blur in the case of unidimensional constructs.

Constructs used for scientific research must have precise and clear definitions that others can use to understand exactly what it means and what it does not mean. For instance, a seemingly simple construct such as *income* may refer to monthly or annual income, before-tax or after-tax income, and personal or family income, and is therefore neither precise nor clear. There are two types of definitions: dictionary definitions and operational definitions. In the more familiar dictionary definition, a construct is often defined in terms of a synonym. For instance, attitude may be defined as a disposition, a feeling, or an affect, and affect in turn is defined as an attitude. Such definitions of a circular nature are not particularly useful in scientific research for elaborating the meaning and content of that construct. Scientific research requires **operational definitions** that define constructs in terms of how they will be empirically measured. For instance, the operational definition of a construct such as *temperature* must specify whether we plan to measure temperature in Celsius, Fahrenheit, or Kelvin scale. A construct such as *income* should be defined in terms of whether we are interested in monthly or annual income, before-tax or after-tax income, and personal or family income. One can imagine that constructs such as *learning*, *personality*, and *intelligence* can be quite hard to define operationally.

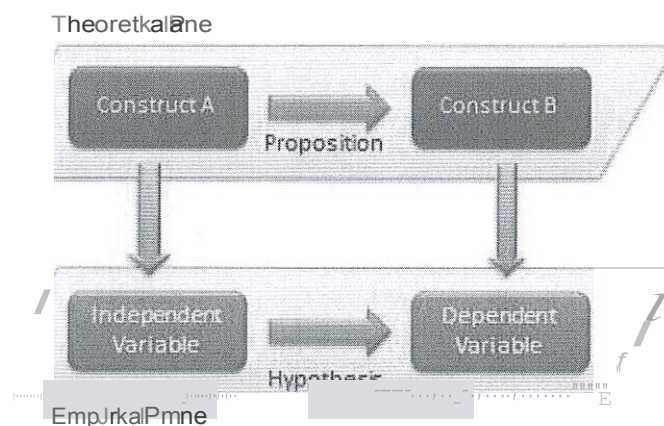


Figure 2.1. The theoretical and empirical planes of research

A term frequently associated with, and sometimes used interchangeably with, a construct is a variable. Etymologically speaking, a variable is a quantity that can vary (e.g., from low to high, negative to positive, etc.), in contrast to constants that do not vary (i.e., remain constant). However, in scientific research, a **variable** is a measurable representation of an abstract construct. As abstract entities, constructs are not directly measurable, and hence, we look for proxy measures called variables. For instance, a person's *intelligence* is often measured as his or her *IQ (intelligence quotient) score*, which is an index generated from an analytical and pattern-matching test administered to people. In this case, *intelligence* is a construct, and *IQ score* is a variable that measures the intelligence construct. Whether IQ scores truly measure one's intelligence is anyone's guess (though many believe that they do), and depending on

whether how well it measures intelligence, the IQ score may be a good or a poor measure of the intelligence construct. As shown in Figure 2.1, scientific research proceeds along two planes: a theoretical plane and an empirical plane. Constructs are conceptualized at the theoretical (abstract) plane, while variables are operationalized and measured at the empirical (observational) plane. Thinking like a researcher implies the ability to move back and forth between these two planes.

Depending on their intended use, variables may be classified as independent, dependent, moderating, mediating, or control variables. Variables that explain other variables are called **independent variables**, those that are explained by other variables are **dependent variables**, those that are explained by independent variables while also explaining dependent variables are **mediating variables** (or intermediate variables), and those that influence the relationship between independent and dependent variables are called **moderating variables**. As an example, if we state that higher intelligence causes improved learning among students, then intelligence is an independent variable and learning is a dependent variable. There may be other extraneous variables that are not pertinent to explaining a given dependent variable, but may have some impact on the dependent variable. These variables must be controlled for in a scientific study, and are therefore called **control variables**.

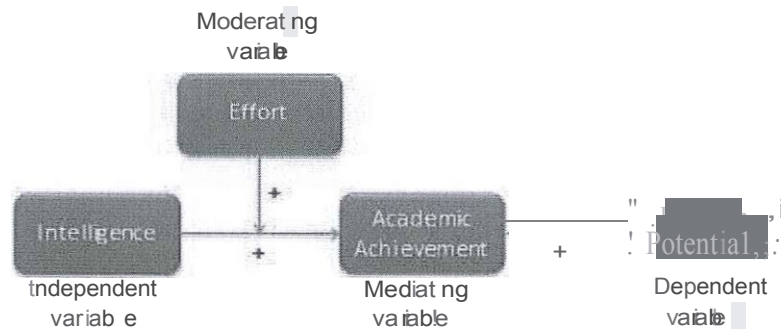


Figure 2.2. A nomological network of constructs

To understand the differences between these different variable types, consider the example shown in Figure 2.2. If we believe that intelligence influences (or explains) students' academic achievement, then a measure of intelligence such as an *IQ score* is an independent variable, while a measure of academic success such as *grade point average* is a dependent variable. If we believe that the effect of intelligence on academic achievement also depends on the effort invested by the student in the learning process (i.e., between two equally intelligent students, the student who puts in more effort achieves higher academic achievement than one who puts in less effort), then *effort* becomes a moderating variable. Incidentally, one may also view effort as an independent variable and intelligence as a moderating variable. If academic achievement is viewed as an intermediate step to higher earning potential, then *earning potential* becomes the dependent variable for the independent variable *academic achievement*, and academic achievement becomes the mediating variable in the relationship between intelligence and earning potential. Hence, variables are defined as an independent, dependent, moderating, or mediating variable based on their nature of association with each other. The overall network of relationships between a set of related constructs is called a **nomological network** (see Figure 2.2). Thinking like a researcher requires not only being able to abstract constructs from observations, but also being able to mentally visualize a nomological network linking these abstract constructs.

Propositions and Hypotheses

Figure 2.2 shows how theoretical constructs such as intelligence, effort, academic achievement, and earning potential are related to each other in a nomological network. Each of these relationships is called a proposition. In seeking explanations to a given phenomenon or behavior, it is not adequate just to identify key concepts and constructs underlying the target phenomenon or behavior. We must also identify and state patterns of relationships between these constructs. Such patterns of relationships are called propositions. A **proposition** is a tentative and conjectural relationship between constructs that is stated in a declarative form. An example of a proposition is: "An increase in student intelligence causes an increase in their academic achievement." This declarative statement does not have to be true, but must be empirically testable using data, so that we can judge whether it is true or false. Propositions are generally derived based on logic (deduction) or empirical observations (induction).

Because propositions are associations between abstract constructs, they cannot be tested directly. Instead, they are tested indirectly by examining the relationship between corresponding measures (variables) of those constructs. The empirical formulation of propositions, stated as relationships between variables, is called **hypotheses** (see Figure 2.1). Since IQ scores and grade point average are operational measures of intelligence and academic achievement respectively, the above proposition can be specified in form of the hypothesis: "An increase in students' IQ score causes an increase in their grade point average." Propositions are specified in the theoretical plane, while hypotheses are specified in the empirical plane. Hence, hypotheses are empirically testable using observed data, and may be rejected if not supported by empirical observations. Of course, the goal of hypothesis testing is to infer whether the corresponding proposition is valid.

Hypotheses can be strong or weak. "Students' IQ scores are related to their academic achievement" is an example of a weak hypothesis, since it indicates neither the directionality of the hypothesis (i.e., whether the relationship is positive or negative), nor its causality (i.e., whether intelligence causes academic achievement or academic achievement causes intelligence). A stronger hypothesis is "students' IQ scores are *positively* related to their academic achievement", which indicates the directionality but not the causality. A still better hypothesis is "students' IQ scores have positive effects on their academic achievement", which specifies both the directionality and the causality (i.e., intelligence causes academic achievement, and not the reverse). The signs in Figure 2.2 indicate the directionality of the respective hypotheses.

Also note that scientific hypotheses should clearly specify independent and dependent variables. In the hypothesis, "students' IQ scores have positive effects on their academic achievement," it is clear that intelligence is the independent variable (the "cause") and academic achievement is the dependent variable (the "effect"). Further, it is also clear that this hypothesis can be evaluated as either true (if higher intelligence leads to higher academic achievement) or false (if higher intelligence has no effect on or leads to lower academic achievement). Later on in this book, we will examine how to empirically test such cause-effect relationships. Statements such as "students are generally intelligent" or "all students can achieve academic success" are not scientific hypotheses because they do not specify independent and dependent variables, nor do they specify a directional relationship that can be evaluated as true or false.

Theories and Models

A **theory** is a set of systematically interrelated constructs and propositions intended to explain and predict a phenomenon or behavior of interest, within certain boundary conditions and assumptions. Essentially, a theory is a systemic collection of related theoretical propositions. While propositions generally connect two or three constructs, theories represent a *system* of multiple constructs and propositions. Hence, theories can be substantially more complex and abstract and of a larger scope than propositions or hypotheses.

I must note here that people not familiar with scientific research often view a theory as a *speculation* or the opposite of *fact*. For instance, people often say that teachers need to be less theoretical and more practical or factual in their classroom teaching. However, practice or fact are not opposites of theory, but in a scientific sense, are essential components needed to test the validity of a theory. A good scientific theory should be well supported using observed facts and should also have practical value, while a poorly defined theory tends to be lacking in these dimensions. Famous organizational research Kurt Lewin once said, "Theory without practice is sterile; practice without theory is blind." Hence, both theory and facts (or practice) are essential for scientific research.

Theories provide explanations of social or natural phenomenon. As emphasized in Chapter 1, these explanations may be good or poor. Hence, there may be good or poor theories. Chapter 3 describes some criteria that can be used to evaluate how good a theory really is. Nevertheless, it is important for researchers to understand that theory is not "truth," there is nothing sacrosanct about any theory, and theories should not be accepted just because they were proposed by someone. In the course of scientific progress, poorer theories are eventually replaced by better theories with higher explanatory power. The essential challenge for researchers is to build better and more comprehensive theories that can explain a target phenomenon better than prior theories.

A term often used in conjunction with theory is a model. A **model** is a representation of all or part of a system that is constructed to study that system (e.g., how the system works or what triggers the system). While a theory tries to explain a phenomenon, a model tries to represent a phenomenon. Models are often used by decision makers to make important decisions based on a given set of inputs. For instance, marketing managers may use models to decide how much money to spend on advertising for different product lines based on parameters such as prior year's advertising expenses, sales, market growth, and competing products. Likewise, weather forecasters can use models to predict future weather patterns based on parameters such as wind speeds, wind direction, temperature, and humidity. While these models are useful, they may not necessarily explain advertising expenditure or weather forecasts. Models may be of different kinds, such as mathematical models, network models, and path models. Models can also be descriptive, predictive, or normative. Descriptive models are frequently used for representing complex systems, for visualizing variables and relationships in such systems. An advertising expenditure model may be a descriptive model. Predictive models (e.g., a regression model) allow forecast of future events. Weather forecasting models are predictive models. Normative models are used to guide our activities along commonly accepted norms or practices. Models may also be static if it represents the state of a system at one point in time, or dynamic, if it represents a system's evolution over time.

The process of theory or model development may involve inductive and deductive reasoning. Recall from Chapter 1 that **deduction** is the process of drawing conclusions about a

phenomenon or behavior based on theoretical or logical reasons and an initial set of premises. As an example, if a certain bank enforces a strict code of ethics for its employees (Premise 1) and Jamie is an employee at that bank (Premise 2), then Jamie can be trusted to follow ethical practices (Conclusion). In deduction, the conclusions must be true if the initial premises and reasons are correct.

In contrast, **induction** is the process of drawing conclusions based on facts or observed evidence. For instance, if a firm spent a lot of money on a promotional campaign (Observation 1), but the sales did not increase (Observation 2), then possibly the promotion campaign was poorly executed (Conclusion). However, there may be rival explanations for poor sales, such as economic recession or the emergence of a competing product or brand or perhaps a supply chain problem. Inductive conclusions are therefore only a hypothesis, and may be disproven. Deductive conclusions generally tend to be stronger than inductive conclusions, but a deductive conclusion based on an incorrect premise is also incorrect.

As shown in Figure 2.3, inductive and deductive reasoning go hand in hand in theory and model building. Induction occurs when we observe a fact and ask, "Why is this happening?" In answering this question, we advance one or more tentative explanations (hypotheses). We then use deduction to narrow down the tentative explanations to the most plausible explanation based on logic and reasonable premises (based on our understanding of the phenomenon under study). Researchers must be able to move back and forth between inductive and deductive reasoning if they are to post extensions or modifications to a given model or theory, or build better ones, which are the essence of scientific research.

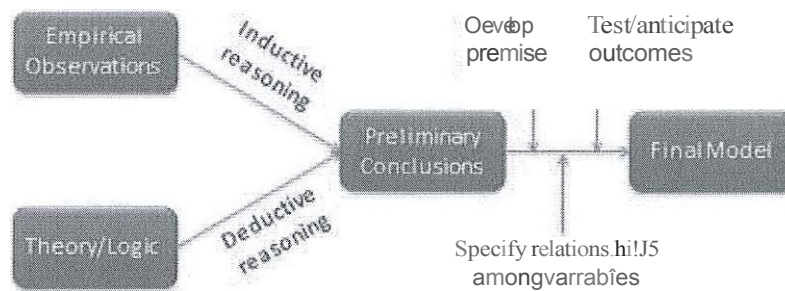


Figure 2.3. The model-building process

INTRoDucTION TO

SOCIAL RESEARCH

QUANTITATIVE & QUALITATIVE APPROACHES

KEITH F PUNCH

ISAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

List of tables	xiii
List of figures	xv
Preface	xvii
Companion website	xix
About the author	xxi
1 Introduction	1
2 Theory and Method in Social Science Research	13
3 Ethics in Social Science Research	35
4 Research Questions	57
5 From Research Questions to Data	73
6 Literature Searching and Reviewing	93
7 Qualitative Research Design	113
8 Collecting Qualitative Data	143
9 Analysing Qualitative Data	167
10 Quantitative Research Design	205
11 Collecting Quantitative Data	227
12 Analysing Quantitative Data	251
13 The Internet and Research	279
14 Mixed Methods and Evaluation	301
15 Research Writing	327
Glossary	347
Appendix 1	355
References	359
Index	377

seven

QUALITATIVE RESEARCH DESIGN

Contents

- / 1 What is research design?
- / Diversity in qualitative research
- J Common themes within the diversity
- 7 1 Case studies
- / The general idea
- 7 Four characteristics of case studies
- J Case studies and generalisability
- / 1 Preparing a case study
- / 4 Ethnography
- 7 4 Introduction
- 1.4 Some main characteristics
- / 4 General comments
- 1.r., Grounded theory
- 7 What is grounded theory?
- / A short history of grounded theory
- / Theory generation versus theory verification
- 1.r., Theoretical sampling: data-collection/data-analysis relationships
- 7. The use of the literature in grounded theory
- / 1 The place of grounded theory research
- 7. Action research
- Chapter summary
- Key terms
- Exercises and study questions
- Further reading
- Notes

Learning Objectives

After studying this chapter you should be able to:

1

Describe the main components of research design, and how questions, design and data are connected

Describe and explain the strategies behind case studies, ethnography, grounded theory and action research

Discuss the strengths and weaknesses of case studies, ethnography, grounded theory and action research

Discuss the potential contribution of case studies, ethnography, grounded theory and action research

Compare and contrast case studies, ethnography, grounded theory and action research, as qualitative research designs

We begin this chapter by looking at research design in general, in order to set a context both for qualitative design in this chapter and quantitative design in Chapter 10. We then focus on four common designs used in qualitative research – case studies, ethnography, grounded theory and action research.

What is research design?

Three uses of the term 'research design' can be distinguished in the literature, roughly ordered from general to specific. *At the most general level* it means all the issues involved in planning and executing a research project – from identifying the problem through to reporting and publishing the results. This is how it is used by Ackoff (1953) and Miller and Salkind (2002), for example. By contrast, *at its most specific level* the design of a study refers to the way a researcher guards against, and tries to rule out, alternative interpretations of results. *Between these two* there is the general idea of design as situating the researcher in the empirical world, and connecting research questions to data (Denzin and Lincoln, 2011). The first view is too broad for our purposes in this chapter, and the second will come up as we go through this chapter and Chapter 10. Here, we will focus on the third use of the term, since we need a way of thinking about design which is general enough to accommodate both qualitative and quantitative approaches.

In this view, research design situates the researcher in the empirical world, and connects the research questions to data, as shown in Figure 7.1. The research design is the basic plan for a piece of research, and includes four main ideas. The first is the strategy. The second is the conceptual framework. The third is the question of who or what will be studied. The fourth concerns the tools and procedures to be used for collecting and analysing empirical materials. Research design thus deals with four main questions, corresponding to these ideas.

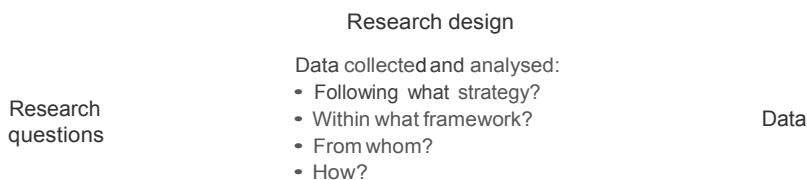


FIGURE 7.1 Research design connects research questions to data

The data will be collected (and analysed):

Following what strategy?
Within what framework?
From whom?
How?

These questions overlap, especially the first two. Also the second question, in particular, is more typical of quantitative designs, although it does apply in some qualitative research. We will now look briefly at each of the four questions.

ii

At the centre of the design of a study is its logic or rationale – the reasoning or the set of ideas by which the study intends to proceed in order to answer its research questions. The term 'strategy' refers to this. Thus, in qualitative research, a multiple case study design involves a strategy (for example: 'the detailed investigation, using multiple sources of data, of a small number of deliberately chosen cases, guided by research questions which focus on comparisons between the cases'). Ethnography and grounded theory are different sorts of strategies the qualitative researcher might use, as is explained in Sections 7.4 and 7.5. Similarly, in quantitative research, the experiment includes a strategy designed to achieve certain comparisons. So does the correlational survey. Answers to the question 'following what strategy?' will differ according to whether the approach is qualitative, quantitative or mixed methods. If qualitative, is the strategy case study, ethnography, grounded theory, action research or some combination of these? If quantitative, is the strategy experimental, quasi-experimental or non-experimental? If there is a combination of quantitative and qualitative approaches, what is the mixture of strategies? Associated with this question of strategy is another important question: To what extent will the researcher manipulate or organise the research situation, as against studying it naturalistically? In other words, to what extent will the researcher intervene in the research situation, contriving it and constructing it for research purposes, as against studying it as it occurs? Qualitative research design is generally non-interventionist. Quantitative research design can vary from extremely interventionist to non-interventionist.

Strategy is important because it drives the design. Or, put another way, behind the design lies a logical rationale for answering the research questions – this is the strategy. In Chapter 14, on mixed methods research (Section 14.4), it is recommended that a short paragraph describing the strategy and design of a study be included in a proposal (and in a dissertation). This same advice applies to all qualitative and quantitative studies.

Framework here means conceptual framework – the conceptual status of the things being studied and their relationship to each other. Quantitative designs typically have well-developed prespecified conceptual frameworks, showing variables and their relationship to each other, whereas qualitative designs show much more variability. While many qualitative studies proceed without a conceptual framework, there is often a role for conceptual frameworks in qualitative research – Miles and Huberman (1994: 18-22) give examples. A conceptual framework may be developed ahead of the study, or it may emerge as the study unfolds. Together with the strategy, it is the conceptual framework that determines how much prespecified structure a study will have.

This question concerns the sampling for the research. In this form, the question is biased towards quantitative studies. The more general question 'Who or what will be studied?' (Denzin and Lincoln, 2011) covers qualitative, quantitative and mixed methods approaches.

This question asks about the tools and procedures to be used in data collection and analysis, topics dealt with in Chapters 8 and 9 for qualitative research and Chapters 11 and 12 for quantitative research.

Together, these four components of research design situate the researcher in the empirical world. Design sits between the research questions and the data, showing how the research questions will be connected to the data, and what tools and procedures to use in answering them. Therefore design needs to follow from the questions and fit in with the data. Design must be driven by strategy. The starting point is the strategy – the logic of the approach by which the data will be used to answer the research questions. Design implements, or formalises, this strategy.

In this book, qualitative and quantitative approaches are both presented under the same three main headings – design, data collection and data analysis. Before considering these headings for qualitative research, the next section looks at the complex nature of this field, stressing its diversity.

In sharp contrast with quantitative research, which seems relatively methodologically unidimensional despite its internal technical debates, a dominant feature of present-day qualitative research is its diversity. Early in their *Handbook*, Oenzin and Lincoln (1994: ix) wrote:

It did not take us long to discover that the 'field' of qualitative research is far from a unified set of principles promulgated by networked groups of scholars. In fact, we have discovered that the field of qualitative research is defined primarily by a series of essential tensions, contradictions and hesitations. These tensions work back and forth among competing definitions and conceptions of the field.

Qualitative research methods is a complex, changing and contested field – a site of multiple methodologies and research practices. 'Qualitative research' therefore is not a single entity, but an umbrella term that encompasses enormous variety.

Four aspects of this diversity concern paradigms, strategies and designs, approaches to data, and methods for the analysis of data. The last three of these are dealt with in this book in Chapters 7, 8 and 9 respectively. This section comments on the diversity of paradigms and perspectives in qualitative research. We need to be aware of the differences between qualitative and quantitative research on this issue.

Paradigm debate and diversity has not been a typical feature of quantitative research. In general, quantitative research has been mainly based on positivism – as Tesch (1990) points out, the whole approach of constructing concepts and measuring variables is inherently positivistic.¹ The situation in qualitative research is quite different, with several different paradigm positions, and much paradigm discussion and debate. By comparison with quantitative research, the field of qualitative research is multidimensional and pluralistic with respect to paradigms. The main alternative paradigms within qualitative research include positivism, post-positivism, critical theory and constructivism, but there are finer distinctions than these and more detailed subdivisions. Furthermore, paradigm developments within qualitative research continue, so that we do not yet have a final picture, although some convergence now seems to be taking place (see Section 2.1). It is important to be aware of this range of paradigm possibilities within qualitative research, especially when reading the literature.

One effect of these developments within qualitative methodology has been to highlight the political nature of much social science research – the recognition that research, like other things people do, is a human construction, framed and presented within a particular set of discourses (and sometimes ideologies), and conducted in a social context with certain sorts of social arrangements, especially involving funding, cognitive authority and power. Both the substantive concepts and the methods research uses are ways of describing the social world for particular purposes, not just abstract and neutral academic tools. In other words, social science research is in part a political process, and always has been. Thus Apple (1991, in Lather, 1994: vii) stresses the inescapably political contexts in which we speak and work, and points out that all of our discourses are politically uninnocent. Or, as Punch (1994) puts

it, politics suffuses all social science research, from the micropolitics of personal relations in a research project, to issues involving research units, universities and university departments, and ultimately government and its agencies.

Some aspects of the political nature and context of research are discussed by Sapsford and Abbott (1996), and by the various writers in *Beyond Methodology* (Fonow and Cook, 1991). A collection of readings edited by Hammersley (1993) considers the politics of research in relation to development studies in the third world, feminism, critical theory, evaluation studies and to methodology and data themselves. Hammersley (1995) also presents a comprehensive review of the changes in the nature of ideas about social research, with reference to political issues and concerns. In Chapter 6 of that book, he undertakes a detailed analysis of the question 'Is social research political?'

Research methods and styles themselves can be seen from this 'politicised' perspective. Sapsford and Abbott (1996) note the argument that choices about research styles are choices that have political elements. Research styles are not neutral, but embody implicit models of what the social world is or should be like, and of what counts as knowledge and how to get it. A consequence of this is that a large area of knowledge is suppressed as 'non-scientific' by the limitations of prevailing research methodologies. Research methods themselves, as a field of study, can be analysed and understood using the approaches and techniques developed within the field to study other things. The politics of research methods, and the university contexts in which choices about methods often occur, are discussed by Jayaratne and Stewart (1991) and by Eisner (1991).

Feminism and *postmodernism* are two perspectives from which the political aspects of research have received a great deal of attention. The former stresses the role of power in research, especially in the traditional hierarchical relationship between researcher and researched. Like critical analysis, and some types of class, race and ethnic studies, feminism also often has emancipation as its goal. The latter perspective often 'foregrounds' power directly, insisting that research is no more immune from the power-knowledge connection than any other human activity (Lather, 1991). Such perspectives apply to virtually every part of the research process – the conception of research itself, the purposes of research, the role of the researcher, approaches to design, data collection and analysis, ethical considerations and evaluative criteria.

on the within the diversity

While qualitative research is much more diverse than quantitative research, there are at the same time important recurrent features in qualitative research.

The first is that a major characteristic of qualitative research, reflected in its strategies and designs, is that it is naturalistic, preferring to study people, things and events in their natural settings. While much quantitative research (for example, an experiment) is not at all naturalistic, quantitative research can be naturalistic also, in studying people in their natural settings, without artificially contriving situations

for research purposes. Some observational studies and correlational surveys fall into this category, but they are likely to have a prefigured conceptual framework and design, with prestructured data. Qualitative designs are more likely to delay conceptualising and structuring of the data until later in the research. They are also much less likely to contrive or create a situation for research purposes.

Beyond this main characteristic, there are several attempts to classify the many varieties of qualitative research by identifying its common features (for example, Tesch, 1990; Wolcott, 1992). A summary of the recurrent elements in qualitative research is given by Miles and Huberman (1994: 6-7) and is reproduced here:

Qualitative research is conducted through an intense and/or prolonged contact with a 'field' or life situation. These situations are typically 'banal' or normal ones, reflective of the everyday life of individuals, groups, societies and organisations.

The researcher's role is to gain a 'holistic' overview of the context under study: its logic, its arrangements, its explicit and implicit rules.

The researcher attempts to capture data on the perceptions of local actors 'from the inside', through a process of deep attentiveness, of empathetic understanding and of suspending or 'bracketing' preconceptions about the topics under discussion.

Reading through these materials, the researcher may isolate certain themes and expressions that can be reviewed with informants, but that should be maintained in their original forms throughout the study.

A main task is to explicate the ways people in particular settings come to understand, account for, take action and otherwise manage their day-to-day situations.

Many interpretations of this material are possible, but some are more compelling for theoretical reasons or on grounds of internal consistency.

Relatively little standardised instrumentation is used at the outset. The researcher is essentially the main 'instrument' in the study.

Most analysis is done with words. The words can be assembled, subclustered, broken into semiotic segments. They can be organised to permit the researcher to contrast, compare, analyse and bestow patterns upon them.

Many of these features will come up in different ways in this and the next two chapters. They provide a good background against which to look at some main qualitative research designs. Against this background, this chapter now describes case studies, ethnographies, grounded theory and action research, as strategies and designs commonly used in qualitative research. There will often be overlap between these four – any particular qualitative study will not necessarily be only one thing or the other. While recognising this, it is still useful to consider each separately.

Case studies

Case studies are now discussed under four headings – the general idea of case studies, some main characteristics, case studies and generalisability, and preparing a case study. Some classic case studies in social science research are shown in Example 7.1.

EXAMPLE 7.1

Examples of case studies

Beachside Comprehensive: A Case Study of Secondary Schooling (Ball, 1981), a study of mixed-ability teaching in a comprehensive school, utilised comparisons of lesson observations between those of the research and those provided by teachers.

Street Corner Society: The Social Structure of an Italian Slum (Whyte, 1955) is a classic example of a descriptive case study. It describes an Italian-American subculture, 'Cornerville', covering one neighbourhood in Boston in the 1940s. Issues of low-income youths and their ability (or inability) to break with neighbourhood ties are discussed.

In Search of Excellence: Lessons from America's Best-Run Companies by Peters and Waterman (1982) is based on more than 60 case studies of large-scale successful American businesses. The text contains cross-case analyses with each chapter dealing with characteristics associated with organisational excellence.

TVA and the Grass Roots: A Study of Politics and Organization, a classic study by Selznick (1949) of the Tennessee Valley Authority (TVA), describes the political behaviour and organisational decentralisation that occurred as a result of the TVA Act. Under this Act the TVA was charged with the duty to plan for the proper use, conservation and development of the natural resources of the Tennessee River drainage basin and its adjoining territory.

The general idea

What is a case study? The basic idea is that one case (or perhaps a small number of cases) will be studied in detail, using whatever methods and data seem appropriate. While there will be specific purposes and research questions, the general objective is to develop as full an understanding of this case as possible. We may be interested only in this case, or we may have in mind not just this case we are studying, but others like it. That raises the question of generalisability, which we will look at later.

In keeping with other approaches in qualitative research, the case study aims to understand the case in depth, and in its natural setting, recognising its complexity and its context. It also has a holistic focus, aiming to preserve and understand the wholeness and unity of the case. Therefore the case study is more a strategy than a method. As Goode and Hatt (1952: 331) pointed out many years ago: 'The case study then is not a specific technique; it is a way of organising social data so as to preserve the unitary character of the social object being studied.' This strategy for

understanding contrasts strongly with the reductionist approach of some quantitative research.

What then is a case? It is difficult to give a full answer to this question, since almost anything can serve as a case, and the case may be simple or complex. But, with Miles, Huberman and Saldana (2013), we can define a case as a phenomenon of some sort occurring in a bounded context. Thus, the case may be an individual, or a role, or a small group, or an organisation, or a community or a nation. It could also be a decision, or a policy, or a process, or an incident or event of some sort, and there are other possibilities as well. Brewer and Hunter (2005) list six types of units that can be studied in research – individuals, attributes of individuals, actions and interactions, residues and artefacts of behaviour, settings, incidents and events, and collectivities. Any of these may be the focus of case study research.

Just as there are different types of cases, there are also different types of case studies. Stake (1994) distinguishes three main types:

the *intrinsic case study*, where the study is undertaken because the researcher wants a better understanding of this particular case;

the *instrumental case study*, where a particular case is examined to give insight into an issue, or to refine a theory; and

the *collective case study*, where the instrumental case study is extended to cover several cases, to learn more about the phenomenon, population or general condition.

The first two of these are single case studies, where the focus is within the case. The third involves multiple cases, where the focus is both within and across cases. It is also called the *multiple case study* or sometimes the *comparative case study*.

Because of the great variation, it is not easy to define the case study. Stake gives a 'pretty loose definition' (1988: 258) – a case study is 'a study of a bounded system, emphasising the unity and wholeness of that system, but confining the attention to those aspects that are relevant to the research problem at the time'. Yin (2013) stresses that a case study is an empirical inquiry that:

investigates a contemporary phenomenon within its real-life context, when the boundaries between phenomenon and context are not clearly evident, and in which multiple sources of evidence are used.

A dictionary of sociological terms defines a case study as:

a method of studying social phenomena through the thorough analysis of an individual case. The case may be a person, a group, an episode, a process, a community, a society, or any other unit of social life. All data relevant to the case are gathered, and all available data are organised in terms of the case. The case study method gives a unitary character to the data being studied by interrelating a variety of facts to a single case. It also provides an opportunity for the intensive analysis of many specific details that are often overlooked with other methods. (Theodorson and Theodorson, 1969)

These definitions highlight four main characteristics of case studies.

1.3.1 Bounded system

The case is a 'bounded system' – it has boundaries. Yin points out that the boundaries between the case and the context are not necessarily clearly evident. Nonetheless, the researcher needs to identify and describe the boundaries of the case as clearly as possible. The case is a case of something. This may seem obvious, but it needs stressing, to give focus to the research, and to make the topic and strategy of the research clear. Identifying what the case is a case of is also important in determining the unit of analysis, an important idea in the analysis of data.

There is an explicit attempt to preserve the wholeness, unity and integrity of the case. The word 'holistic' is often used in this connection. At the same time, since not everything can be studied, even about one case, specific focus and within-case sampling are required. Research questions help to define this focus.

Multiple sources of data and multiple data collection methods are very likely to be used, typically in a naturalistic setting. Many case studies will use sociological and anthropological field methods, such as observations in natural settings, interviews and narrative reports. But they may also use questionnaires and numerical data. This means that the case study is not necessarily a totally qualitative technique, though most case studies are predominantly qualitative.

1.3.2 Case studies and generalisability

A common criticism of the case study concerns its generalisability: 'This study is based on only one case, so how can we generalise?' Because this reaction is so common, we need to take this question seriously.

The first point is to ask whether we would want to generalise from a particular case study. There are two types of case study situations where generalisation would not be the objective. *First*, the case may be so important, interesting or misunderstood that it deserves study in its own right. Or it may be unique in some very important respects, and therefore worthy of study. These are examples of Stake's intrinsic case study. It is not the intention of such a study to generalise, but rather to understand this case in its complexity and its entirety, as well as in its context. *Second*, a strong argument can often be made about studying the 'negative case'. This is where a particular case seems to be markedly different from the general pattern of other cases, perhaps even completely opposite to them, creating the need to understand why this case is so different. The logic is that we can learn about the typical by studying the atypical, as when we study disease in order to learn about health. This is Stake's second type of case study, the instrumental case study. Therefore, whether a case study should even seek to generalise, and claim to be representative, depends on the context and purposes of the particular project. Generalisation should not necessarily be the objective of all research projects, whether case studies or not (Denzin, 1983).

Aside from these two situations, however, there are many case studies where we do have in mind more than just the case being studied, and where we do want to find something more broadly applicable. How can a case study produce something that might be generalisable? There are two main ways that a case study can produce potentially generalisable results. Both depend on the purposes of the case study, and especially on the way its data are analysed. *The first is by conceptualising*, the *second is by developing propositions*. In both instances, the findings from a case study can be put forward as being potentially applicable to other cases.

To *conceptualise* means that, on the basis of the disciplined and in-depth study of this case, and using methods for the analysis of data that focus on conceptualising rather than on describing (for example, those described in Chapter 9 under grounded theory analysis), the researcher develops one or more new concepts to explain some aspect of what has been studied. Indeed, to develop such new concepts may require the sort of in-depth study that is only possible in a case study. To *develop propositions* means that, based on the case being studied, the researcher puts forward one or more propositions – they could be called hypotheses – about concepts or elements or factors within the case. These can then be assessed for their applicability and transferability to other situations. This turns the traditional model of research around. In traditional quantitative research, we often begin with propositions or hypotheses – they are inputs into the research. In this view of case study research, we end with them – they become outputs of the research.

In neither of these instances will the one case study have proved the generalisability of its findings. But it can certainly suggest such generalisability, putting forward concepts or propositions for testing in further research. Clearly, every case that can be studied is in some respects unique. But every case is also, in some respects, similar to other cases. The question is whether we want to focus on what is unique about a particular case, or on what is common with other cases. At different times we need to do each of these, and we need to be aware of when we are doing each. This is a matter to be addressed in the purposes and research questions that are developed to guide a case study. When generalisability is a goal, and we are focusing on the potential common elements in a case, it is necessary for the analysis of the case study data to be conducted at a sufficient level of abstraction. The more abstract a concept, the more generalisable it is. Developing abstract concepts and propositions raises the analysis above simple description, and in this way a case study can contribute potentially generalisable findings.

The generalisation process is not mechanical, though this is more freely recognised in qualitative research than in quantitative research. There have been some attempts to see the complexity of generalisation in the quantitative context (for example, Bracht and Glass, 1968), but it is still widely regarded there as generalisation from a sample to a population. In fact, however, as Firestone (1993) points out, there are three levels of generalisation – generalisation from sample to population, analytic or theory-connected generalisation, and case-to-case transfer. Similarly, Stake (1988: 260) distinguishes between scientific generalisation, arrived at by experimentation and induction, and naturalistic generalisation,

where general understandings are furthered by case studies and experience in individual events.²

While on this lack-of-generalisability criticism of case study research, which is often a 'knee jerk' reaction to the case study, we should note the central role given to the case method of teaching in professional schools of business, medicine and law, as well as nursing, public administration, social work and psychoanalysis (Reinharz, 1992). In these training situations, historical cases are studied in great detail and are used to train managers, doctors, lawyers, and so on, to deal with situations they will encounter in the future. This clearly underlines the potential generalisability of knowledge built from case studies. If every case were totally unique, there would be no transferability of knowledge from one case to another, and little point in the case method of training.

Case studies have had an ambiguous place in social science research (Reinharz, 1992), and historically there has often been a disapproving attitude towards the case study. This attitude is usually based on the generalisability criticism and is expressed in the condescending remark 'that's only a case study'. This book takes a different view. Properly conducted case studies, especially in situations where our knowledge is shallow, fragmentary, incomplete or non-existent, have a valuable contribution to make in social science research, in three main ways:

The *first* is what we can learn from the study of a particular case, in its own right. As noted, the case being studied might be unusual, unique or not yet understood, so that building an in-depth understanding of the case is valuable. This might cover all of the three types of case study described by Stake.

Second, only the in-depth case study can provide understanding of the important aspects of a new or persistently problematic research area. This is particularly true when complex social behaviour is involved, as is the case in much social science research. Discovering the important features, developing an understanding of them and conceptualising them for further study, is often best achieved through the case study strategy. Following this line of argument, it may be that too much research has tried to go straight to measurement and quantitative mapping, without a fuller understanding of the phenomena and processes involved that are best achieved by case studies.

Third, the case study can make an important contribution in combination with other research approaches.

For example, a case study ahead of a survey can give direction to *this* survey not otherwise possible without the understanding built from the case study. Similarly, a survey could be followed by, or done in conjunction with, one or more case studies. Because of the limitations of the survey, the case study can 'flesh out' the picture in a way that is both crucial to our understanding, and not possible using more superficial techniques. In addition, the case study may be particularly appropriate in a student project or dissertation, where there are limited resources, including time.

These potential contributions of the case study counter the disapproving attitude described above. At the same time, this critical attitude can have validity,

es ecially when a case study is standing alone, not integrated with other approaches tots subject matter and simply descriptive, or when more is claimed from its findings than the data can bear. Therefore, because of these criticisms, and because of the diversity within case study research, it seems especially important to be clear on the rationale behind the case study and on its purpose(s). That means clarifying the strategy of the case study and developing research questions to guide the study, either ahead of it or as focal points in the case become clear.

7.1.4 Preparing a case study

We can now summarise what has been said into a set of guidelines for preparing a case study. A case study research proposal would need to:

- be clear on what the case is and on what it is a case of, in a way that anticipates and connects to the strategy behind the research;
- be clear on the need for the study of this case and on the general purpose(s) of this case study;
- translate this general purpose into specific purposes and research questions (these may emerge during the early empirical work);
- identify the overall strategy of the case study, especially whether it is one case or multiple cases, and why;
- show how the strategy leads to the case(s) selected for study;
- show what data will be collected, from whom and how;
- show how the data will be analysed.

The last point will come up again, in Chapter 9 especially when we look at levels of abstraction in the analysis of qualitative data. Similarly, the first point, on identifying and bounding the case, has implications for the unit of analysis in the study and for the analysis of the study's data.

Ethnography

This section has three parts. First, it summarises the introduction to ethnography given by Hammersley and Atkinson in their well-known textbook on the subject. Second, it identifies some important features of the ethnographic approach to research. Third, it makes some general comments about the place of ethnography in social science research. Examples of ethnographic studies are shown in Example 7.2. The term ethnography itself comes from cultural anthropology. 'Ethno' means people or folk, while 'graphy' refers to describing something. Thus ethnography means describing a culture and understanding a way of life from the point of view of its participants – ethnography is the art and science of describing a group or culture (Fetterman, 2010; Neuman, 1994). Fielding (2008) discusses the origins of ethnography and surveys the history of its use in British colonial and American research.

Hammersley and Atkinson (2007) take a 'fairly liberal' view of ethnography, whereby the ethnographer participates, overtly or covertly, in people's daily lives for an extended period of time, watching what happens, listening to what is said asking questions and collecting any other relevant data. They point out ethnography's connection to naturalism, a way of doing social research developed by ethnographers in the face of the difficulties they saw with positivism. In naturalistic research, unlike other approaches, the social world is studied as far as possible in its natural state, undisturbed by the researcher. Research uses methods that are sensitive to the nature of the setting, and the primary aim is to describe what happens in the setting, and how the people involved see their own actions, others' actions and the context.

Drawing especially on symbolic interactionism (see Box 7.1), but also on phenomenology and hermeneutics, naturalism sees social phenomena as quite different in character from physical phenomena. The basic ideas here are that human behaviour is based upon meanings that people attribute to and bring to situations, and that behaviour is not 'caused' in any mechanical way, but is continually constructed and reconstructed on the basis of people's interpretations of the situations they are in.

Symbolic Interactionism

There is a natural affinity between ethnography and symbolic interactionism. But symbolic interactionism is also of great general importance in qualitative research, beyond ethnography. Symbolic interactionism is a general theory about human behaviour which stresses that people define, interpret and give meaning to situations, and then behave in response to these definitions, interpretations and meanings. It is the 'actor's definition of the situation', or the insider's view, which is important in accounting for human behaviour, not some 'objective' reality of the situation itself. The insider's view and the meanings of situations and actions to the participants are paramount, and symbolic interactionist researchers want access to this view and these meanings. Theoretical treatments of symbolic interactionism are given by Blumer (1969) and Woods (1992). (Examples of the use of symbolic interactionism in education research can be found in van den Berg, 2002; Evans, 2007; and O'Donoghue, 2007.)

Therefore, to understand behaviour, we need an approach that gives access to the meanings that guide behaviour. It is the capacities we have all developed as social actors -the capacity to do participant observation (see Chapter 8) -which can give

5 this access. As participant observers we can learn the culture or subculture of the people we are studying, and learn to understand the world as they do. Classic anthropological studies demonstrate how this approach is used to study societies other than our own, but it can be used for the study of all societies, including our own. This is because there are many different layers of cultural knowledge within any society, especially modern industrialised society.

Thus ethnography:

exhibits the capacity that any social actor possesses for learning new cultures, and the objectivity to which this process gives rise. Even where he or she is researching a familiar group or setting, the participant observer is required to treat this as anthropologically strange, in an effort to make explicit the presuppositions he or she takes for granted as a culture member. In this way, it is hoped, the culture is turned into an object available for study. Naturalism proposes that through marginality, in social position and perspective, it is possible to construct an account of the culture under investigation that both understands it from within and captures it as external to, and independent of, the researcher: in other words, as a natural phenomenon. Thus, the *description* of cultures becomes the primary goal. (Hammersley and Atkinson, 2007: 9)

The concept of culture is central in ethnography. *Culture* can be thought of as a shared set of meanings or a cognitive map of meanings (Spradley, 1980). The cultural knowledge that any group of people have is their knowledge of this map. Ethnography has developed within anthropology as the central strategy to study culture, and many anthropologists consider cultural interpretation to be ethnography's main contribution. A full discussion of the concept of culture is beyond our scope here, but useful references are Keesing (1976), Haviland et al. (2013) and Howard (1997). Derived from culture, the concept of subculture has great applicability in social science research. Any stable group of people develops over time a shared set of meanings, and in this way a subculture develops. Drawing on this, research can study ethnographically the subculture of any stable group, whether children or adults.

We can summarise this introduction to ethnography using the words of a prominent educational ethnographer:

Ethnography means, literally, a picture of the way of life of some identifiable group of people. Conceivably, those people could be any culture-bearing group, in any time and place. In times past, the group was usually a small, intact, essentially self-sufficient social unit, and it was always a group notably 'strange' to the observer. The anthropologist's purpose as ethnographer was to learn about, record, and ultimately portray the culture of this other group. Anthropologists always study human behaviour in terms of cultural context. Particular individuals, customs, institutions, or events are of anthropological interest as they relate to a generalised description of the life-way of a socially interacting group. Yet culture itself is always an abstraction, regardless of whether one is referring to culture in general or to the culture of a specific social group. (Wolcott, 1988: 188)

The overarching characteristic of the ethnographic approach is its *commitment to cultural interpretation*. The point of ethnography is to study and understand the cultural and symbolic aspects of behaviour and the context of this behaviour, whatever the specific focus of the research. This specific focus is typically either some group of people, or a case (or a small number of cases), focusing on culturally significant behaviour. In addition to this central characteristic, we can identify six important and interrelated features of the ethnographic approach.

When studying a group of people, ethnography starts from the assumption that the *shared cultural meanings of the group* are crucial to understanding its behaviour. This is part of its commitment to cultural interpretation. As Goffman (1961: ix-x) says: 'any group of persons – prisoners, primitives, pilots or patients – develop a life of their own that becomes meaningful, reasonable and normal once you get close to it. . . .' The ethnographer's task is to uncover that meaning.

The ethnographer is sensitive to the *meanings* that behaviour, actions, events and contexts have, in the eyes of the people involved. What is needed is the *insider's perspective* on those events, actions and contexts. As Spindler and Spindler (1992: 73) point out: 'Sociocultural knowledge held by social participants makes social behaviour and communication sensible. Therefore a major part of the ethnographic task is to elicit that knowledge from informant participants.' The ethnographic study will be designed, and its data collection techniques organised, in line with this.

The group or case will be studied in its *natural setting*. A true ethnography therefore involves the researcher becoming part of that natural setting (Fielding, 2008). This explains why participant observation, discussed in Chapter 8, is the favoured method in ethnographic research. To understand any group, or any culturally significant act, event or process, it is necessary to study behaviour in its natural setting, with special reference to the symbolic world associated with this behaviour.

An ethnography is likely to be an *unfolding and evolving sort of study*, rather than a prestructured one. As part of developing a focus for the study, it will not normally be clear what to study in depth until some fieldwork has been done. While specific research questions and perhaps hypotheses will be used in the research, they are more likely to develop as the study proceeds, rather than to be formulated ahead of the research. This point also applies to data collection procedures. Data collection in ethnography may use several techniques, but any structuring of the data, or of data collection instruments, will be generated in situ, as the study unfolds.

From the point of view of *data collection techniques*, ethnography is *eclectic*, not restricted. Any techniques might be used, but fieldwork is always central. An ethnographic fieldwork continuum would range from direct non-participant observation to participant observation, then to ethnographic interviewing with one or more informants, and then to the words of the people themselves (often called, in ethnographic writing, the 'voices of the natives'). Data collection may well range across this whole continuum in an ethnography, and it may be further supplemented by anything that gives a fuller picture of the live data, such as film or audio records, documents, diaries, and so on. It may also use structured and quantitative questionnaires, with scaled variables, though these would be developed as the study proceeds.

6 Ethnographic *data collection* will typically be *prolonged and repetitive*. There is both a general and a specific reason for this. The general reason is that the reality being studied, the meanings, symbolic significance and cultural interpretation, exists on several levels. It takes time for a researcher to gain access to the deeper and most important levels of this reality (Woods, 1992). The specific reason is that the ethnographic record needs to be comprehensive and detailed, and typically focuses on things that happen again and again. The ethnographer therefore needs to observe this a sufficient number of times. Closure is achieved by recognising the point at which nothing new about its cultural significance is being learned.

7 : ,,,.....

While ethnography is a distinctive strategy, there is no one design for an ethnographic study. Its design may overlap, in whole or in part, with other designs. Thus, for example, it may use elements of the case study or grounded theory approaches, which are consistent with its orientation. It can also be used in combination with field experimentation and with surveys. Whatever the specific design, ethnography typically uses relatively unstructured empirical materials, a small number of cases and a style of analysis and writing that stresses description and interpretation (Atkinson and Hammersley, 1994). Ethnography is also both process and product. 'Process' means that it is a particular approach to research and has a particular distinctive way of going about it. 'Product' means that a certain type of research report (sometimes called the ethnographic record or a full ethnographic description) will be produced. The term 'an ethnography' illustrates the idea of ethnography as a product.

A full-scale ethnography means carrying out a detailed and demanding study, with fieldwork and data collection running over a long period of time. Where these demands exceed the time and resources of one project, there is nonetheless great value in bringing the ethnographic approach to the topic. Thus elements of the ethnographic approach, or 'borrowing ethnographic techniques' (Wolcott, 1988), are used in some social science research projects, rather than producing full-scale ethnographies. Borrowing from ethnographies is also helpful in qualitative social science research through the study of subcultures, as noted.

When would the ethnographic approach be most appropriate? In general, when we need to understand the cultural context of behaviour, and the symbolic meaning and significance of the behaviour within this context. The ethnographic approach, being a method of discovery, is particularly useful when we are dealing with something new, different or unknown. It is an excellent way of gaining insight into a culture, sub-culture or social process, particularly those in complex behavioural settings, and particularly those involving other cultures and subcultures, including those of the organisations and institutions of the modern world. The ethnographic approach can sensitise us to the cultural context and symbolic significance of behaviour we need to understand, in a way that other research approaches cannot. As Fielding (2008: 265) points out, it is often pathbreaking, and, 'as a means of gaining a first insight into a culture or social process, as a source of hypotheses for detailed investigation using other methods, it is unparalleled'.

With the culture and subculture of different groups, and of different institutions and organisations, there is both ample scope and an important contribution for the ethnographic approach in social science research. Some prominent ethnographic studies are shown in Example 7.2.

—EXAMPLE 7.2—

Ethnographies

Translated Woman: Crossing the Border with Esperanza's Story (Behar, 1993) is the Life story of a Mexican Indian woman who was reputed to have bewitched her former husband for abusing her and leaving her for another woman. Rumours of her witchcraft powers were reinforced when her husband suddenly went blind.

When Prophecy Fails: A Social and Psychological Study of a Modern Group that Predicted the Destruction of the World, a participant observation study by Festinger et al. (1964), was carried out opportunistically with two small groups who claimed to have received messages from a planet, 'Clarion', predicting a catastrophic flood in three months. The researchers and some hired observers joined the group and conducted intensive investigations before the predicted disaster and afterwards during the period of disconfirmation.

The National Front (Fielding, 1981) is an ethnography of an extreme right racist organisation. The researcher joined the group as a member and conducted participant observation at meetings and interviews with party officials and opponents of the party, as well as content analysis of party documents.

McLaren's (1986) ethnographic study, *Schooling as a Ritual Performance: Towards a Political Economy of Educational Symbols and Gestures*, is of an inner-city Catholic school in Toronto, Canada, where the school population is largely made up of Portuguese and Italian students. McLaren analyses body postures and gestures of students and generates a theoretical framework for conceptualising embodied meaning and power.

The Man in the Principal's Office: An Ethnography is Wolcott's (1973) inquiry into the behaviour of one elementary school principal. The researcher spent two years following a typical school principal in all of his professional and many of his private activities.

Grounded theory

As a research strategy, grounded theory is specific and different. At the same time it cuts across the other strategies and designs discussed in this chapter, and is 'currently the most widely used and popular qualitative research method across a wide range of disciplines and subject areas' (Bryant and Charmaz, 2007a: 1). This book has two sections on grounded theory, one in this chapter and one in Chapter 9. This is because grounded theory is both a strategy for research and a way of analysing data. Chapter 9 (Section 9.5) deals with

rounded theory analysis. In this chapter we deal with grounded theory as a strategy under six headings:

- What is grounded theory?
- A short history of grounded theory
- Theory generation research versus theory verification research
- Theoretical sampling/data-collection/data-analysis relationships
- The use of the literature in grounded theory
- The place of grounded theory research

Examples of grounded theory studies are shown below in Example 7.3, and more are noted in Chapter 9.

EXAMPLE 7.3

Examples of grounded theory studies

Using a database of 33 interviews with academic department chairpersons, Creswell and Brown (1992) in 'How chairpersons enhance faculty research: a grounded theory study' developed a grounded theory relating categories of chair influence to faculty scholarly performance.

Fresh Starts: Men and Women after Divorce (Cauhape, 1983) describes the processes by which men and women rebuild their social worlds after mid-life divorce. Participants were upwardly mobile professional men and women, who were originally from non-professional backgrounds.

Awareness of Dying (Glaser and Strauss, 1965) was the first publication reporting the original grounded theory studies. Those studies (and this book) focus on the process of dying: what happens when people die in hospitals, how hospitals manage the situation, and the interaction between staff and patients. The research was carried out at six hospitals in San Francisco.

Time for Dying (Glaser and Strauss, 1968) was the second report of the grounded theory study. This book is based on intensive fieldwork combining observation and interviewing in the six hospitals. The focus again is on the organization of terminal care in hospitals, and the aim in the book is to describe the temporal features of dying, seeing dying itself as a social process.

From Practice to Grounded Theory (Chenitz and Swanson, 1986: Chapters 14 to 19) describes six grounded theory studies dealing with topics such as 'Getting around with emphysema' and 'Entry into a nursing home as status passage'.

The focus in Davis's (1973) study *Living with Multiple Sclerosis: A Social Psychological Analysis* was on patients with multiple sclerosis who, in certain circumstances, took the initiative in furthering the continuity of their care.

What is grounded theory?

The first point to make is that grounded theory is not a theory at all. It is a research strategy, or, from some points of view, a research approach or method.

Grounded theory is a research strategy whose purpose is to generate theory from data. 'Grounded' means that the theory will be generated on the basis of data; the theory will therefore be grounded in data. 'Theory' means that the objective of collecting and analysing the research data is to generate theory to explain the data. The essential idea in grounded theory is that explanatory theory will be developed inductively from data. Grounded theory, then, is an overall strategy for doing research. To implement this strategy, grounded theory has a particular set of techniques and procedures. As well as the grounded theory strategy, we can therefore talk also about grounded theory analysis – that style of analysis which uses procedures to develop a theory grounded in the data, as described in Chapter 9.

Short history of grounded theory

A brief look at the history of grounded theory helps in understanding it, and in seeing its present place in social science research. Its early history can be traced primarily through five key publications. In the 1960s, Glaser and Strauss began collaborative work in medical sociology, and published two landmark studies of dying in hospitals (Glaser and Strauss, 1965, 1968). These books had an important impact, and represented a different style of empirically based sociology. In response to numerous 'how did you do it?' requests from readers after *Awareness of Dying* was published, the authors wrote a book that detailed the methods they had developed and used in the dying studies. This book, published in 1967 under the title of *The Discovery of Grounded Theory*, was the first description of the method and the first key publication about grounded theory. According to Strauss and Corbin (2008: 326), *The Discovery of Grounded Theory* had three purposes – to offer a rationale for theory that was grounded, to suggest the logic for and specifics of grounded theories, and to legitimate careful qualitative research. In the years after its publication, first Glaser and then Strauss taught a grounded theory-style seminar in qualitative analysis at the University of California in San Francisco.

While a good deal of research using grounded theory to investigate a variety of phenomena was published by numerous graduates of this programme, the next methodological work, and the second key publication, came 11 years later with Glaser's *Theoretical Sensitivity*, published in 1978. Its purposes were to update methodological developments in grounded theory and to help analysts develop theoretical sensitivity. Once again, while studies reporting grounded theory research continued to be published, it was another nine years before the next methodological statement. This was Strauss's *Qualitative Analysis for Social Scientists*, published in 1987, and the third key grounded theory publication. In this book the focus is broadened to qualitative analysis in general, but grounded theory still plays the central role. It is described as 'a handbook of sorts for the better understanding of social phenomena through a particular style of qualitative analysis of data (*grounded theory*)'. That mode of doing analysis . . . is designed especially for *generating and testing theory*' (p. x emphasis in original).

The fourth key publication came in 1990, with Strauss and Corbin's *Basics of Qualitative Research*, subtitled 'Grounded Theory Procedures and Techniques'. It is addressed to researchers in various disciplines who aim to build theory through the analysis of qualitative data. It presents the analytic mode of grounded theory, and stresses that skill in this method of analysis is learnable by anyone who takes the trouble to study its procedures. This provoked, in response, the fifth key publication – Glaser's critique of the Strauss and Corbin book – titled *Basics of Grounded Theory Analysis* subtitled 'Emergence vs Forcing' (Glaser, 1992). In this book Glaser sets out to correct what he takes to be the misconceptions about grounded theory evident in the Strauss and Corbin book.

These five publications give the early history of the development of grounded theory. They are not the only methodological statements on grounded theory from that period, but they are the main ones. Since the early 1990s, however, there has been considerable further development and diversification of grounded theory approaches and methods. Main recent features include constructivist grounded theory (Charmaz, 2006) and the 2007 publication *The Sage Handbook of Grounded Theory* (Bryant and Charmaz, 2007b). As Bryant and Charmaz point out in Chapter 1 of the *Handbook*, grounded theory methods now seem to have taken on a life of their own. A basic three-way classification within the present-day diversification of grounded theory would include: (a) 'traditional' or 'classical' grounded theory, as practised by Glaser and his followers, (b) followers of the Strauss and Corbin approach, and (c) followers of Charmaz's constructivist grounded theory. On the other hand, on a more detailed level, Denzin identifies seven versions. Thus grounded theory is best viewed today not as one method, but as a family of methods (Bryant and Charmaz, 2007a: 10).

2.3 Theory generation versus theory verification

Grounded theory has as its explicit purpose the generation of theory from data. This raises the contrast between research that aims to generate theory and research that aims to verify theory. As pointed out in Chapter 2, this contrast represents a difference in research styles. Traditionally, much research, especially quantitative research, has followed the theory verification model, as indicated in the importance it has traditionally given to the role of the hypothesis. Many research methods texts insisted that hypotheses were central to research and that, since the hypothesis was deduced from some more general theory, the point of the research was the testing of theory.

As noted in Chapter 4, this book takes a different view of the hypothesis, recommending that it be included only when appropriate. In the grounded theory approach, which aims to generate theory, no 'up-front' theory is proposed, and no hypotheses are formulated for testing ahead of the research. The research does not start with a theory from which it deduces hypotheses for testing. It starts with some research questions and an open mind, then it moves to data, aiming to end up with

a theory. This emphasis was developed deliberately by Glaser and Strauss as a reaction to the exclusive insistence on theory verification research, especially in the American sociology of the 1950s.

It is useful to make this theory generation-vs-verification contrast sharply, in order to highlight the difference in research styles. But in fact, in practice, the distinction is not so sharp. For while we may start without a theory, and have the objective of creating one, it is not long into the theorising process before we are also wanting to test theoretical ideas that are emerging. So, in fact, theory generation depends on progressive verification, as well. Another way of saying this is that grounded theory is essentially an inductive technique, but it uses deduction as well. It stresses induction as the main tool for theory development, but, in developing the theory, deduction will also often be necessary.

7.5.4 Theoretical sampling: data-collection/ data-analysis relationships

Grounded theory has a specific approach to this topic, which is different from many other approaches. (It is not unique, however – see Hughes (1958) and Becker (1971).)

In the traditional view of research, data collection is a discrete stage in the research, usually to be completed before data analysis begins. In grounded theory, the pattern is different. Guided by some initial research questions, the researcher will collect a first set of data, often quite small. At this point, analysis of the data begins, using the procedures to be described in Chapter 9. The second set of data will be collected after the first analysis of data, guided by emerging directions in this analysis. This is the principle of theoretical sampling – the idea that subsequent data collection should be guided by theoretical developments that emerge in the analysis of previously collected data. This cycle of alternation between data collection and analysis will not stop at two repetitions. It continues until theoretical saturation is achieved – that is, until new data are not showing new theoretical elements, but are rather confirming what has already been found. This pattern is shown in Figure 7.2.

It is becoming more common to find this sort of data-collection/data-analysis relationship in qualitative research today. It is different from traditional research, but it resembles what we normally do in everyday life, when we encounter a puzzling situation. Like much else in grounded theory, it models the way humans have always learned. In this respect, grounded theory is faithful to its philosophical roots in pragmatism (Glaser and Strauss, 1967).

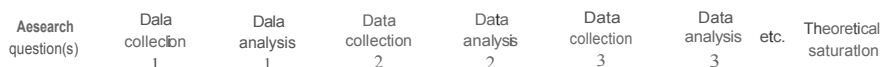


FIGURE 7.2 Theoretical sampling: data-collection/data-analysis relationships

Grounded theory also has a different perspective on this matter from other research approaches. The difference lies in how the literature is dealt with, and when it is introduced, and follows from the stress that grounded theory places on theory generation.

If a satisfactory theory already exists on a particular topic, there is no point in mounting a study to generate a new theory about the topic. The rationale for doing a grounded theory study is that we have no satisfactory theory on the topic, and that we do not understand enough about it to begin theorising. In this case, we will want to approach the data as open-mindedly as possible, guided by research questions. While a general comment on the literature may be necessary to orient a study, and to show the lack of satisfactory theory, the problem with a detailed substantive review of the literature in advance of a study is that it can strongly influence us when we begin working with the data.

As is detailed in Chapter 9 we want to begin the analysis by finding categories and concepts within the data, not by bringing them to the data, from the literature or from anywhere else. In such a case, it makes sense to delay the literature-reviewing stage of the work, at least until conceptual directions within the data have become clear. We will introduce the literature later than would normally be done, seeing the relevant literature as further data for the study. This is the key concept in using the literature in grounded theory – the literature is seen as further data to be fed into the analysis, but at a stage in the data analysis when theoretical directions have become clear. This use of the literature is consistent with the overall logic of grounded theory research. The whole approach is organised around the principle that theory that is developed will be grounded in data.

7.5.6 The place of grounded theory research

It is not surprising that grounded theory has become a widely used approach in qualitative research. I think there are five main reasons for this:

- 1 While much is said in the research methodology literature about the need to generate theory in research, very little is said about *how* to do this. Grounded theory explicitly addresses this question.

It represents a coordinated, systematic but flexible overall research strategy, in contrast to the ad hoc and uncoordinated approaches that have sometimes characterised qualitative research.

- 2 It brings a disciplined and organised approach to the analysis of qualitative data. In the qualitative research context, with its history of a lack of well-formulated methods for the analysis of data, this point has great appeal.

There are impressive demonstrations of what the grounded theory approach can produce in a research area. These began with the dying studies of Glaser and Strauss, and have continued, initially in the area of medical sociology, and now much more broadly (Bryman and Charmaz, 2007b).

A fifth reason has to do with the identification of research problems from professional practice, and from organisational and institutional contexts. In these situations, a traditional hypothesis-testing approach is not appropriate. Many of these problems confronting social science researchers, especially in applied areas, are substantively new, because they come from new developments in professional practice and/or from newly developing organisational contexts. Empirical research, much of it qualitative, is needed in these areas, and the theory verification approach would be inappropriate. The theory generation approach of grounded theory has much to recommend it in these substantively new areas, where there is a lack of grounded concepts for describing and explaining what goes on. Grounded theory appeals because it concentrates on discovering concepts, hypotheses and theories.

+: on research

Early in *The Handbook of Action Research*, editors Reason and Bradbury (2007: 1) tell us that there is no short answer to the question 'What is action research?' Rather, the term is used for a family of related strategies that share certain important common ideas, while differing in details of their approach to the research. The differences have led to a variety of names by which such researchers describe their approach - technical action research, practical action research, emancipatory action research, participatory action research and collaborative action research are examples, along with feminist action research - but the generic term action research probably encompasses most of the approaches (Kemmis and McTaggart, 2000: 567). This section concentrates on the main common ideas behind the different strands of action research.³

The central idea is conveyed by the term 'action research' itself. Action and research are brought together: action researchers 'engage in careful, diligent inquiry, not for purposes of discovering new facts or revising accepted laws or theories, but to acquire information having practical application to the solution of specific problems related to their work' (Stringer, 2004: 3). Action research brings together the acting (or the doing) and the researching (or the inquiring). In contrast to the ideas of inquiry for its own sake and building knowledge for its own sake, action research aims to design inquiry and build knowledge for use in the service of action to solve practical problems. Therefore, in action research, the inquiry deliberately starts from a specific practical or applied problem or question. Its whole purpose is to lead to action to solve this practical problem or answer this practical question. As Reason and Bradbury (2008: 1) say, action research 'seeks to bring together action and reflection, theory and practice, in participation with others, in the pursuit of practical solutions to issues of pressing concern to people'. And again (2008: 2): 'A primary purpose of action research is to produce practical knowledge that is useful to people in the everyday conduct of their lives.' In a similar vein, Stringer's five-part action research sequence shows 'basic research' in four parts (research design, data gathering, data analysis, communication), with action research adding a fifth part - action itself - to these.

Stringer's five-part action research sequence shows clearly that research itself is central in the sequence. That is, systematic, disciplined inquiry – research – is brought to bear on a practical problem that requires a solution – action. All of this is done in a carefully organised framework. This systematic, disciplined inquiry – this research – is, of course, empirical. Therefore it draws on the approaches to research covered in this book. Thus action research may involve quantitative data methods and designs, qualitative data methods and designs, or mixed methods data and designs. While action research is usually thought of as a qualitative approach, and is included here under qualitative research designs, it does not rely only on qualitative data. On the contrary, it uses quantitative data whenever they are appropriate and available. It is like case study research in this respect.

An important characteristic of action research, which sets it apart from other designs, is that it is usually *cyclical* in nature, reflecting the fact that people usually work towards solutions to their problems in cyclical, iterative ways. The words 'cycle', 'spiral' and (less often) 'helix' are used by writers on action research to describe this. They convey the idea that the one piece of research leading to the one set of actions is not the end of the process, but rather the start of a cycle or spiral. The research produces outcomes that lead to the taking of action, but this in turn generates further questions for research, which in turn generates further action, and so on. Kemmis and McTaggart (2000: 595-6) diagram the action research spiral, and write that, while difficult to describe the process as a series of steps, participatory action research is generally thought to involve a spiral of self-reflective cycles of:

- planning a change;
- acting and observing the consequences of the change;
- reflecting on these processes and consequences, and then;
- replanning;
- acting and observing;
- reflecting, and so on.

Stringer begins with the action research cycle, then broadens this to the action research helix and then spiral. Whichever version we consider, the main idea here is that action research is repetitive, continuing and cyclical.

For many people, the spiral of cycles of self-reflection, involving planning, acting and observing, reflecting, replanning and so on, has become the dominant feature of action research as an approach. For Kemmis and McTaggart, however, there are seven additional important features of participatory action research – it is a social process, participatory, practical and collaborative, emancipatory, critical, recursive, and it aims to transform both theory and practice.

Just as action research does not separate inquiring from doing, neither does it separate the researcher from the researched. An older version of action research, especially in education research in the 1970s, located the two roles in the one person – the teacher became the action researcher. This led to credibility problems for action research, since most teachers did not have the research skills to communicate effectively to an often-sceptical research community. Now the action and the



research are seen as different roles, and are typically done by different people, but collaboration and participation between the different people are stressed. Stringer (2004) distinguishes practitioner research in education from action research in education on this very point. When the teacher steps back, reflects, collects information, observes classroom interaction and so on, this is practitioner research. When the teacher engages others in the process of inquiry, with the intent of solving an educational work problem together, this is action research. Collaborative participation becomes central.

Similarly, Kemmis and McTaggart (2000: 595) believe that, while some action research depends on solitary processes of systematic self-reflection by the action researcher, the steps in the self-reflection spiral are best undertaken collaboratively by co-participants in the research process. This is why they prefer the term *participatory action research*. Their formulation highlights the role of participation and collaboration in some types of action research. When participation and collaboration are involved, action research develops new research relationships, and often works towards building a community of learners. Whether this happens or not, the researcher and the researched become co-researchers, collaborating participants in the action research.

Action research has diverse origins. Many writers trace it back to the social experiments of Kurt Lewin in the 1940s, but Reason and Bradbury (2007: 2--4) identify other important influences as well. These include the contemporary critique of positivist science and scientism, Marxism ('the important thing is not to understand the world but to change it'), the liberating perspectives on gender and race, the practices of experiential learning and psychotherapy, and some types of spiritual practices. Kemmis and McTaggart (2000: 568) also note the connection of participatory research to liberation theology and Third World movements aimed at social transformation. In education, action research became popular in the 1970s, but then declined in popularity and credibility in the 1980s, only to re-surface strongly in the 1990s. An indication of its present popularity in education research is the vast literature on action research in education (Stringer, 2004). An indication of its present prominence in social science research in general comes from the recent and already mentioned *Handbook of Action Research* (Reason and Bradbury, 2007).

Chapter summary

Research design connects research questions to data. It is based on a strategy, often involves a conceptual framework, and shows from whom, and how, data will be collected and analysed.

Multiple paradigms, perspectives and strategies and designs characterise present-day qualitative research. At the same time, there are important common features across this diversity. In case study research, one case (or a small number of cases) is studied in depth, in context, in its natural setting and holistically. There should be a logic behind case selection, and research questions and multiple sources of data are normally involved.

Ethnography focuses on the way of life of some group of people, which can only be understood from the insider's perspective. Culture – as a shared set of meanings – is the central concept, and multiple sources of data, mostly qualitative, are used by the ethnographer to uncover cultural meanings.

Grounded theory is a research strategy whose objective is to generate explanatory theory grounded in data. It has evolved today into a family of methods, with distinctive concepts and approaches.

Action research is a family of related approaches which stress the bringing together of action and research, in a cyclical pattern directed at solving practical problems, often in a participative situation.

KEY TERMS

Research design: connects research questions to data; design is based on a strategy, and shows from whom, and how, data will be collected and analysed

case study: the detailed, holistic and in-context study of one case or a small number of cases

Ethnography: a research strategy which focuses on uncovering the shared meanings which develop among any stable group of people

Culture: the set of meanings shared by a group of people, without which their behaviour and actions cannot be understood

Symbolic interactionism: a general theory which stresses that people behave in terms of the way they define (or interpret, or give meaning to) situations

Insider's perspective: the definition, interpretation or meaning given to a situation by the participants in that situation

Grounded theory: a research strategy for generating theory grounded in data

Theoretical sampling: later stages of data collection are guided by theoretical developments emerging from earlier data

Action research: a research strategy which combines action and research in cyclical spirals to focus on the solution to a problem

Exercises and study questions

List four questions that can help us understand research design. What is the function of research design?

What is meant by research strategy, and what is its relationship to research design?

What is a case study, and what are its strengths and weaknesses as a research strategy?

Outline the strategy and design for the study of a case (an individual, a group, an organisation, a decision, etc.) with which you are familiar. Follow the points given in Section 7.3.4.

What does ethnography mean? What is its connection to anthropology, and to the concept of culture?

How can ethnography be applied in social science research?

What is meant in research by the insider's perspective?

Why did Glaser and Strauss use the term 'grounded' to describe the grounded theory method they developed?

What does it mean to say that grounded theory is best seen as a family of methods?

What is theoretical sampling?

What key characteristics of action research make it a distinct research strategy?

Further reading

- Ragin, C.C. and Becker, H.S. (eds) (1992) *What is a Case? Exploring the Foundations of Social Inquiry*. New York: Cambridge University Press.
- Stake, R.E. (1988) 'Case study methods in educational research: seeking sweet water', in R.M. Jaeger (ed.), *Complementary Methods for Research in Education*. Washington, DC: American Educational Research Association. pp. 253-300.
- Stake, R.E. (1994) 'Case studies', in N.K. Denzin and Y.S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: SAGE. pp. 236-47.
- Stake, R.E. (2006) *Multiple Case Study Analysis*. New York: Guilford Press.
- Yin, R.K. (2013) *Case Study Research: Design and Methods*. 5th edn. Thousand Oaks, CA: SAGE.
- Agar, M. (1986) *Speaking of Ethnography*. Beverly Hills, CA: SAGE.
- Atkinson, P. and Hammersley, M. (1994) 'Ethnography and participant observation', in N.K. Denzin and Y.S. Lincoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: SAGE. pp. 248-61.
- Atkinson, P., Delamont, S., Coffey, A., Lofland, J. and Lofland, L. (eds) (2007) *Handbook of Ethnography*. London: SAGE.
- Crang, M. and Cook, I. (2007) *Doing Ethnographies*. London: SAGE.
- Fetterman, D.M. (2010) *Ethnography Step by Step*. 3rd edn. Thousand Oaks, CA: SAGE.
- Gobo, G. (2007) *Doing Ethnography*. London: SAGE.
- Hammersley, M. and Atkinson, P. (2007) *Ethnography: Principles in Practice*. 3rd edn. London: Routledge.
- Spindler, G. and Spindler, L. (1992) 'Cultural process and ethnography: an anthropological perspective', in M.D. LeCompte, W.L. Milroy and J. Preissle (eds), *The Handbook of Qualitative Research in Education*. San Diego, CA: Academic Press. pp. 53-92.
- Wolcott, H.F. (1988) 'Ethnographic research in education', in R.M. Jaeger (ed.), *Complementary Methods for Research in Education*. Washington, DC: American Educational Research Association. pp. 187-249.
- Woods, P.H. (1986) *Inside Schools: Ethnography in Educational Research*. London: Routledge & Kegan Paul.

- aryant, A. and Charmaz, K. (eds) (2007) *The Sage Handbook of Grounded Theory*. Thousand Oaks, CA: SAGE.
- Charmaz, K. (2006) *Constructing Grounded Theory*. Thousand Oaks, CA: SAGE.
- Glaser, B. (1978) *Theoretical Sensitivity*. Mill Valley, CA: Sociology Press.
- Glaser, a. (1992) *Basics of Grounded Theory Analysis: Emergence vs Forcing*. Mill Valley, CA: sociology Press.
- Glaser, B. and Strauss, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- strauss, A. (1987) *Qualitative Analysis for Social Scientists*. New York: Cambridge University Press.
- strauss, A. and Corbin, I. (1994) 'Grounded theory methodology: an overview', in N.K. Denzin and v.s. Uncoln (eds), *Handbook of Qualitative Research*. Thousand Oaks, CA: SAGE. pp. 273-85.
- strauss, A. and Corbin, I. (2008) *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. 3rd edn. Thousand Oaks, CA: SAGE.

- Herr, K. and Anderson, G.L. (2005) *The Action Research Dissertation*. London: SAGE.
- Kemmis, S. and McTaggart, R. (2000) 'Participatory action research', in N.K. Denzin and Y.S. Lincoln (eds), (2011), *Handbook of Qualitative Research*. 4th edn. Thousand Oaks, CA: SAGE.
- Reason, P. and Bradbury, H. (eds) (2007) *Handbook of Action Research*. 2nd edn. London: SAGE.
- Sagor, R. (2004) *Action Research Guidebook*. London: SAGE.
- stringer, E. (1996) *Action Research: A Handbook for Practitioners*. Thousand Oaks, CA: SAGE.
- Stringer, E. (2007) *Action Research in Education*. 2nd edn. Upper Saddle River, NJ: Pearson.
- Taylor, C., Wike, M. and Baser, J. (2006) *Doing Action Research*. London: SAGE.

Notes

We should be careful, however, about labelling all quantitative research positivistic, for two reasons. One is that the term 'positivism' has many different interpretations (Blaikie, 1993); the other is that some researchers (for example, Marsh, 1982) point out that some quantitative work is not positivist.

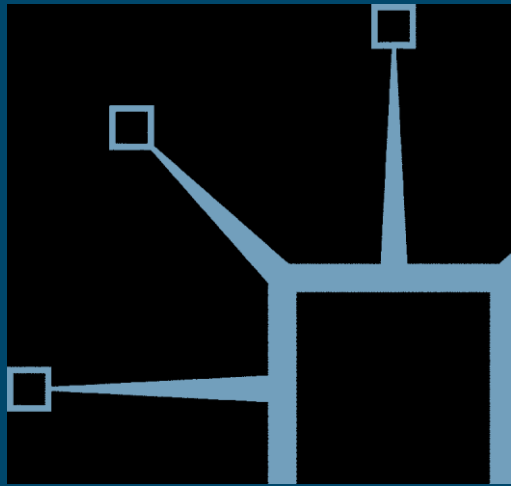
- J Stake also reports a personal communication from Julian Stanley: 'When I want to find out something important for myself, I often use the case study approach' (1988: 262). This statement is worth bringing to the attention of critics of case study research, coming as it does from a respected quantitative researcher, and a major contributor to its literature. Kemmis and McTaggart (2000: 568-72) identify seven approaches within the general area of participatory action research. They are: participatory research, critical action research, classroom action research, action learning, action science, soft systems approaches and industrial action research.

Qualitative Methods in International Relations

A Pluralist Guide

Edited by

Audie Klotz and Deepa Prakash



Qualitative Methods in International

Research Methods Series

General Editors: **Bernhard Kittel**, Professor of Social Science Methodology, Department of Social Sciences, Carl von Ossietzky Universität Oldenburg, Germany and **Benoît Rihoux**, Professor of Political Science, Université catholique de Louvain (UCL), Belgium

In association with the European Consortium for Political Research (ECPR), Palgrave Macmillan is delighted to announce the launch of a new book series dedicated to producing cutting-edge titles in Research Methods. While political science currently tends to import methods developed in neighbouring disciplines, the series contributes to developing a methodological apparatus focusing on those methods which are appropriate in dealing with the specific research problems of the discipline.

The series provides students and scholars with state-of-the-art scholarship on methodology, methods and techniques. It comprises innovative and intellectually rigorous monographs and edited collections which bridge schools of thought and cross the boundaries of conventional approaches. The series covers both empirical-analytical and interpretive approaches, micro and macro studies, and quantitative and qualitative methods.

Titles include:

Audie Klotz and Deepa Prakash (*editors*)

QUALITATIVE METHODS IN INTERNATIONAL RELATIONS

A Pluralist Guide

Forthcoming titles include:

Lane Kenworthy and Alexander Hicks (*editors*)

METHOD AND SUBSTANCE IN MACRO-COMPARATIVE ANALYSIS

Research Methods Series

Series Standing Order ISBN 978-0230-20679-3-hardcover

Series Standing Order ISBN 978-0230-20680-9-paperback

(outside North America only)

You can receive future titles in this series as they are published by placing a standing order. Please contact your bookseller or, in case of difficulty, write to us at the address below with your name and address, the title of the series and one of the ISBNs quoted above.

Customer Services Department, Macmillan Distribution Ltd, Houndmills, Basingstoke, Hampshire RG21 6XS, England

Qualitative Methods in International Relations

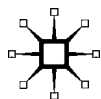
A Pluralist Guide

Edited by

Audie Klotz and Deepa Prakash

*Department of Political Science, The Maxwell School of Citizenship
and Public Affairs, Syracuse University, USA*

palgrave
macmillan



Editorial matter, selection, Introduction and Conclusion © Audie Klotz and Deepa Prakash 2008

All remaining chapters © respective authors 2008

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No paragraph of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, 90 Tottenham Court Road, London W1T 4LP.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The authors have asserted their rights to be identified as the authors of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2008 by
PALGRAVE MACMILLAN

Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010

Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN-13: 978-0-230-54239-6 hardback

ISBN-10: 0-230-54239-5 hardback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Qualitative methods in international relations: a pluralist guide/
edited by Audie Klotz and Deepa Prakash.

p. cm.

Includes bibliographical references and index.

ISBN 0-230-54239-5 (alk. paper)

1. International relations--Research--Methodology.

2. International relations--Study and teaching.

I. Klotz, Audie, 1962- II. Prakash, Deepa, 1978-

JZ1234.Q83 2008

327.072--dc22

2007050062

10 9 8 7 6 5 4 3 2 1

17 16 15 14 13 12 11 10 09 08

Printed and bound in Great Britain by

CPI Antony Rowe, Chippenham and Eastbourne

Contents

<i>List of Tables</i>	vii
<i>List of Figures</i>	viii
<i>Acknowledgments</i>	ix
<i>Notes on Contributors</i>	xi
1 Introduction <i>Audie Klotz</i>	1
Part I: Research Design	
2 Thinking Tools <i>Anna Leander</i>	11
3 Feminist Methodological Reflection <i>Brooke Ackerly</i>	28
4 Case Selection <i>Audie Klotz</i>	43
Part II: Classic Qualitative Tools	
5 Discourse Analysis <i>Iver B. Neumann</i>	61
6 Historical Representations <i>Kevin C. Dunn</i>	78
7 Ethnographic Research <i>Hugh Gusterson</i>	93
8 Process Tracing <i>Jeffrey T. Checkel</i>	114
Part III: Boundary Crossing Techniques	
9 Political Personality Profiling <i>Jerrold M. Post</i>	131

10	Content Analysis <i>Margaret G. Hermann</i>	151
11	Pragmatic Analysis <i>Gavan Duffy</i>	168
12	Agent-based Modeling <i>Matthew J. Hoffmann</i>	187
Part IV: Implications		
13	'Qualitative' Methods? <i>Samuel Barkin</i>	211
14	Practicing Pluralism <i>Deepa Prakash</i>	221
	<i>Bibliography</i>	238
	<i>Index</i>	255

List of Tables

11.1	Interaction between Henry Kissinger (HAK), Mao Zedong (MAO), and Zhou En Lai (ZHOU) (February 1973, Beijing)	175
11.2	Analysis of italicized proposition in Table 11.1	178
11.3	Hypothetical illustration of Rescher's Formalism	180
11.4	Argument analysis of the Kissinger-Mao interaction	182
11.5	Model of Mao's reasoning	183

List of Figures

11.1	Flow chart of pragmatic analysis procedures	170
A.1	Population predictions – low noise, no entrepreneur 10 agents, 6% noise	203
A.2	Population predictions – high noise, no entrepreneur 10 agents, 10% noise	203
A.3	Population predictions – low noise, no entrepreneur present 10 agents, 6% noise	204
A.4	Population predictions – high noise, no entrepreneur present 10 agents, 10% noise	205

Acknowledgments

Three groups of people deserve our heartfelt thanks for their influence on this project. Without them, there would be no book.

The initial idea derived from discussions between Audie Klotz and Cecelia Lynch related to their co-authored *Strategies for Research in Constructivist International Relations*. That book focused on the toolbox. Since it could not include everything, here we concentrate on the tools. While Cecelia did not contribute a chapter of her own, she greatly influenced our emphasis on pluralism. Delivering the tools, of course, would not have been possible without the contributors, all of whom responded enthusiastically to our invitation to write about their use of methods. They probably never imagined how demanding their editors would be! We appreciate their willingness to rethink and rewrite, sometimes more than once.

A number of people at the Maxwell School of Syracuse University influenced the particular form of the project. At Peg Hermann's suggestion, and with her invaluable financial support through the Moynihan Institute of Global Affairs, most of the contributors presented their first drafts to a live student audience. The Moynihan staff provided logistical assistance for those visits. Peg also generously hosted a seaside dinner in San Diego during the 2006 annual meeting of the International Studies Association, the only opportunity for most of the authors to talk in person. Gavan Duffy originally created PSC 694, Qualitative Political Analysis, and he graciously let us commandeer it for 2 years. Azamat Sakiev provided research assistance at the earliest stage, and both his and Deepa Prakash's work were funded by the Political Science Department.

Most books are written by professors, who supposedly know best how to teach their subject matter. In contrast, learners played a central role in this project. We especially thank all student commentators in the Fall 2005 cohort, who provided detailed suggestions and prodded our visiting experts in insightful ways. Revised versions had to pass muster with a second group of students in Fall 2006, although they did not have the fun of grilling the authors in person. In addition, students in both years experimented with these guidelines in their individual homework

assignments, often in very creative ways. The editors tried to channel all of this feedback, but any credit for making the chapters accessible goes to these students. We are grateful for their willingness to contribute to work in progress and to embrace learning as a process.

A.K. and D.P.
Syracuse, NY

Notes on Contributors

Brooke Ackerly, Associate Professor at Vanderbilt University, is co-editor of *Feminist Methods in International Relations*, among other work bridging feminist political theory and international relations.

Samuel Barkin, Associate Professor at the University of Florida, has published on a wide range of topics in international relations, including international political economy, institutions, environmental politics, and theory.

Jeffrey T. Checkel is Professor of Political Science at the University of Oslo and Adjunct Research Professor at the International Peace Research Institute in Oslo. Among his publications is the edited collection *International Institutions and Socialization in Europe*, and he serves as co-editor of the *Sage Series on the Foundations of International Relations*.

Gavan Duffy is Associate Professor at Syracuse University, where he investigates political conflict and conducts formal and computational linguistic analyses of strategic political interactions.

Kevin C. Dunn, Associate Professor at Hobart and William Smith Colleges, works at the intersection of international relations and African politics. He is author of *Imagining the Congo* and co-editor of *African Guerrillas, Identity and Global Politics*, and *Africa's Challenge to International Relations Theory*.

Hugh Gusterson, Professor of Cultural Studies at George Mason University, offers an anthropological perspective on science and technology issues in international affairs, including two books on nuclear scientists, *Nuclear Rights* and *People of the Bomb*. He also co-edited *Cultures of Insecurity* and *Why America's Top Pundits are Wrong*.

Margaret G. Hermann is Gerald B. and Daphna Cramer Professor of Global Affairs and Director of the Daniel Patrick Moynihan Institute of Global Affairs at the Maxwell School of Syracuse University. She is a widely published expert on comparative foreign policy, political leadership, and political psychology.

Matthew J. Hoffmann is Assistant Professor at the University of Toronto. His research interests include global environmental governance, social constructivism, complexity theory, and agent-based modeling. His book *Ozone Depletion and Climate Change* explored social norm dynamics in environmental politics with both computational simulations and process tracing case studies.

Audie Klotz, Associate Professor at Syracuse University, is the co-author of *Strategies for Research in Constructivist International Relations*, among other publications. Her current research on state identity compares immigration policies in South Africa, Canada, and Australia.

Anna Leander, Associate Professor in the Department of Intercultural Communication and Management at the Copenhagen Business School, works on issues of private authority in international relations, focusing on the use of force.

Iver B. Neumann is Professor of Russian Studies at Oslo University and Research Professor at the Norwegian Institute of International Affairs. He has used discourse analysis in studies of diplomacy and popular culture, including his co-edited book, *Harry Potter in International Relations*.

Jerrold M. Post, Professor in the Elliott School of International Affairs at George Washington University and Director of its Political Psychology program, has developed personality profiling as a framework for both scholarly and applied analysis of political leaders.

Deepa Prakash is a PhD candidate in Political Science at Syracuse University. Her research interests are in critical security studies, the discourse on 'terrorism,' and the identities of non-state actors, with a focus on South Asia.

1

Introduction

Audie Klotz

Debates across the social sciences rely on philosophical markers, notably the contemporary polarization between the so-called ‘positivists’ and ‘post-modernists.’ These labels are contested. Few ‘positivists’ rely on a narrow definition of falsification, and many ‘post-modernists’ reject extreme relativism. But the division is also grounded in some legitimate ontological and epistemological differences. For instance, positivists resist including language as a form of observable behavior, and those who reject by assumption the salience of culture or language need not debate how best to study meanings. Post-modernists, in turn, generally see concerns over rigorous analysis as a hallmark of a putatively flawed scientific approach to human action. One unfortunate result of this pervasive divide is a limited appreciation of the insights offered by scholars working within alternative frameworks. It leaves little common ground for analyzing the role of rhetoric in foreign policy choice, for instance.

Despite their abstract nature, the main terrain of these disputes is the realm of empirical research, including the delineation of legitimate research questions, allocation of funding for projects, and employment in the profession. For example, the conflation of ideas with ideology in the traditional ‘Realist’ characterization of ‘Idealism,’ still dominant in the field of International Relations (IR), privileges materialist explanations. The epistemological question of interpretation gets sidelined, because ideas are assumed not to matter as much as military capabilities. As a result, IR privileges a certain form of diplomatic history that lacks serious consideration of discourse analysis. And that can make it harder for scholars employing post-modern inspired approaches to get published in mainstream journals or get jobs at research universities (particularly in the United States).

2 Introduction

Much has been written about this situation (see Hall 1999 on the philosophical issues and Steinmetz 2005 on the disciplinary ones). It has even spawned a 'perestroika' movement in Political Science aimed at opening up that discipline (Monroe 2005). But we still lack true intellectual engagement. Discussion remains an abstract positioning at the level of ontology or epistemology. Yet researchers need practical answers at the level of methodology: How *should* scholars interpret meanings? In IR, recent literature provides plenty of useful illustrations (such as the diverse contributions in Katzenstein 1996 and Weldes *et al.* 1999) but little about the practical trade-offs between techniques for analyzing language. What is at stake in selecting from discourse, speech acts, and semiotics – or even content analysis? When might it be justifiable to combine tools drawn from different analytical traditions – can discourse analysis or semiotics inform the construction of a dictionary for context-sensitive computerized coding, for instance?

We think that refocusing on methodological questions can break down the insularity of scholarly communities, because the justification for practical choices in empirical research exposes underlying ontological and epistemological assumptions (Klotz and Lynch 2007). We concentrate on IR (broadly defined) to provide a degree of empirical overlap. This helps to reveal how researchers wrestle with similar sorts of decisions that require the translation of abstract assumptions into concrete practices. Why do researchers define key concepts differently? How much 'data' is enough? What makes one interpretation better than another? We may still disagree on procedures and standards, but dialogue over methodology forces us to state the goals of our research, clearly define our core concepts, and set out our theoretical assumptions. Then, if warranted, researchers can expand their tools, or at least be able to understand a broader range of relevant literatures.

Many advocates of pluralism already seek to bridge the qualitative-quantitative split through the use of mixed methods. Statistical analysis can certainly be combined with case studies to capture causality in terms of conditions and mechanisms. Yet the presumption remains that positivism and post-modernism are incompatible. For instance, Sprinz and Wolinsky-Nahmias actively promote pluralism, including formal models, but (mis-) characterize post-modernism as lacking methodology (2004: 5). Consequently, we have no guidelines for determining when post-modern analytical techniques are *similar*, *complementary*, or *incompatible* with prevailing positivist approaches. For instance, both rational choice and literary criticism offer theoretical templates for

historical narrative, but these remain *very* distinct literatures. Pluralism, as currently practiced, falls short.

Our starting point is 'qualitative' methods, because of the absence of sufficient guidelines for applying these tools. In contrast, courses in a wide array of statistical techniques are readily available. This gap creates a misperception that historiography and ethnography, for example, do not need to be taught to students and that experienced scholars intuitively know how to use interviews or textual analysis. Researchers of all generations continue to share tales of frustration about learning the trade through trial and error.

An increase in qualitative methods books and courses across the social sciences recognizes this need for practical lessons (for a sampling, see the syllabi posted on the website of the Consortium for Qualitative Research Methods hosted at Arizona State University, that many publishers are expanding their offerings in this area is readily evident in their current catalogues). Those written by political scientists remain oriented toward testing theories and making causal arguments (King *et al.* 1994; Brady and Collier 2004; George and Bennett 2005; Goertz 2006; Trachtenberg 2006; Gerring 2007). Most ignore post-modernism or reject it explicitly; a few offer asides about limited compatibility. Two notable exceptions lean the other way, in defense of critical theory and interpretation: the compendium by Ackerly *et al.* (2006) of feminist approaches in IR and the commentaries compiled by Yanow and Schwartz-Shea (2006).

We adopt a broader view, taking seriously the goals of *both* post-modernist and positivist researchers. This book starts from the assumption that 'qualitative' methods are somehow linked to meaning. But we leave open the boundaries of what should be labeled qualitative, as well as the possibilities for combining qualitative with quantitative and formal approaches. The chapters in this book present a cross-sample of perspectives, ranging from interpretation inspired by Foucault to mechanism-seeking process tracing all the way to agent-based modeling. While the authors work within the field of IR (or international studies, as some might prefer), they bring the insights of other fields, opening up an interdisciplinary conversation.

The contributors offer detailed guidance on how to apply specific tools of analysis and how to circumvent some inherent limitations. All are accomplished scholars who share, with extraordinary candor, their successes and failures. Since fostering use of a broader range of analytical tools requires breaking down the barriers constructed by epistemological polarization, we also asked them to consider whether it would be

appropriate (and if so, when) to combine their primary tools with other qualitative, quantitative, and/or formal techniques.

Part I segues from ontology and epistemology to methodology via research design. Any project is grounded in particular literatures, and the theories contained in those literatures provide a specific vocabulary to characterize the empirical world. Theories, by their nature, simplify and privilege certain aspects of that world. Yet few works on methodology help aspiring researchers get from those ontological assumptions, manifest in theories and concepts, to methodological choices. Illustrating with applications of Pierre Bourdieu's field analysis, Anna Leander, in Chapter 2, offers four steps for translating key concepts into empirical work: asking questions, exploring the relationship between key concepts, figuring out how to apply those concepts, and reflecting on the ways in which those concepts, in turn, can create social realities.

Extending Leander's comments on reflexivity, Brooke Ackerly, in Chapter 3, points out that some concepts, notably gender, embed scholars in their own social environments, presenting researchers with a series of potential dilemmas in the design of their studies. Tensions start with the formulation of key questions and range from very practical issues of sampling to the ethical implications of publishing. For those striving to sensitize themselves to inequalities, in both theoretical formulations and research practices, she offers 'curb cutting' as a pedagogical tool that trains people to view the world through different interpretive lenses.

Leander's and Ackerly's shared emphasis on context and interpretation are, for many, the hallmarks of 'qualitative' case-based research. But in Chapter 4, Audie Klotz uses their insights to challenge the common treatment of case studies as a 'method.' Case selection, she argues, is part of research design, and a variety of methods can be used to analyze them. Researchers should, therefore, clarify their questions, their concepts, and their logic of comparison before tackling the two tasks specific to case selection: defining a 'case' of something and mapping out the universe of *possible* cases (including non-cases). She then assesses three strategies: single cases, paired comparisons, and the elusive category of 'More-than-Two but Not-a-Lot.'

Especially for the Classic Qualitative Tools covered in Part II, we selected authors who would draw on examples from IR because researchers in our field lack teaching materials that address the particularities we face. Discourse analysis by a literary theorist, for instance, may operate at an aesthetic level that does not capture politics or policy concerns. In contrast, Iver Neumann, in Chapter 5, suggests ways to

turn censorship into an analytical advantage, among other insights. He translates the meta-theory of discourse into four methodological steps. The first is a precondition: a degree of cultural competence. From there, he guides readers through the delimitation of texts and subsequent mapping of the representations that comprise discourse. The final step is to untangle the layering of dominant and subordinate discourses.

Diplomatic history's narrative approach has long dominated qualitative analysis in IR (even after the 'history' versus 'science' debates of the 1960s) and is amply represented in the burgeoning methods literature (Elman and Elman 2001; Trachtenberg 2006). Alternatively, offering a post-modern perspective in Chapter 6, Kevin Dunn shifts down from Neumann's macro-historical level to explore agency in the creation of representations and contestation over them. After clearly situating his work ontologically and epistemologically, including its differences from causal analysis, he offers concrete advice on tracking down archival materials destroyed by arsonists and coping with the overwhelming amount of textual, visual, official, popular, and other materials appropriate for his genealogical approach to history.

Unlike historiography, ethnography appears infrequently as a tool of analysis in IR, perhaps because advice from an anthropologist working in a rural village is of limited use to someone seeking to do participant observation in a government department. But anthropology as a field is shifting away from the local in isolation, and as Hugh Gusterson demonstrates in Chapter 7, participant observation and interviewing can indeed help to answer questions about international security. Security clearance may be a distinctive barrier, but access to any field site presents challenges. Starting, like Neumann and Dunn, from a theoretical perspective informed by Foucault, Gusterson presents ethnography as a tool for mapping meanings, but he carries this out at the micro-level of individuals within their communities.

Given the penchant for qualitative analysis in IR to focus on individuals as key actors in historical narratives, Jeffrey Checkel's use of process tracing, in Chapter 8, presents an extension of a traditional approach, rather than an alternative one. By linking process tracing to the study of causal mechanisms generally, and by illustrating with independent and dependent variables beyond the foreign policy arena narrowly defined, he opens up possibilities for its application at diverse levels of analysis and across fields of study. Checkel also discusses some practical considerations of using elite interviews, official documents, and secondary sources to distinguish various dynamics of decision-making

and collective identity formation as micro-level mechanisms of socialization.

Part III continues this focus on individuals and micro-level analysis, albeit in radically different ways. Each in its own way challenges what typically would be considered a qualitative method yet still captures some element of its hallmarks: meaning, interpretation, and context. Therefore, we call these Boundary Crossing Techniques, because they force researchers to reconsider what, if any, characteristics should define qualitative research.

In Chapter 9, Jerrold Post brings psychology and psychiatry to the task of figuring out what makes leaders 'tick.' His technique of Political Personality Profiling is a variant of the single case study, one which draws on personal history and comparison via personality type. No special training is required, he points out, only a sensitivity to psychologically minded types of observations that enable the researcher to identify the characteristics and patterns that clinicians use for diagnosis. More generally, his approach offers one answer to questions about how deeply analysts can delve into the minds of their research subjects.

Margaret Hermann, in Chapter 10, asks many of the same research questions about political leadership as does Post (and as many analysts of foreign policy do), but she uses Content Analysis as her tool for analyzing individuals at a distance. She delineates eight generic steps that any researcher should think through in order to analyze large quantities of textual (and visual) materials. Along the way, she challenges some of the myths that many interpretive scholars have about this approach; current software programs, for example, do allow for context-sensitive coding.

Gavan Duffy remains skeptical about getting into the minds of these leaders, even at a distance, and offers an alternative approach in Chapter 11 that concentrates on communications between individuals (such as foreign policy makers). Influenced by Anglo-American speech act theory, he applies formal logic to texts in order to create replicable interpretations. Pragmatic Analysis contrasts with the post-modern inspired approaches of Neumann, Dunn, and Gusterson, as Duffy holds out the future possibility of using computers to provide systematic analyses of discourse.

Taking formalization one step further, Matthew Hoffmann makes a case, in Chapter 12, for adding agent-based models to the interpreter's toolkit. He argues that this particular form of computer simulation can capture key dynamics of mutual constitution. Yet he insists that all

models are heuristics, and, therefore, remain the basis for interpretation rather than objective analysis.

Part IV steps back from particular tools to Implications for pluralism in research and teaching. In Chapter 13, Samuel Barkin returns to the broad debates alluded to in this introduction. He remains skeptical of the term ‘qualitative’ and cautions against any naïve embrace of pluralism. More optimistically, Deepa Prakash in Chapter 14 highlights teaching tools that work especially well. She draws on her own experiences and those of her peers as they experimented with the guidelines offered in the manuscript versions of this book, as well as her perusal of assignments described in other syllabi. Together, these two chapters give scholars plenty of ideas for teaching and learning without falling into the trap of reifying the category of qualitative methods.

This book provides both an introduction to unfamiliar techniques and a guide for better application of familiar tools. Those designing a course might want to assign the chapters in order, while someone looking primarily to use a particular approach can safely skip to that section. Others may wish to concentrate on particular themes, clustering the chapters that focus on textual analysis, for instance, or perhaps those concerned with individuals as agents. Cross-references within each chapter provide suggestions for identifying such threads. Readers trying to figure out how to combine various techniques would benefit from reading the Research Design and Implications sections before delving into the toolbox. While controversies in contemporary IR and Political Science instigated the creation of this book, we hope that these chapters – in whatever order they are read – will prove useful to researchers seeking to practice pluralism across the social sciences.

This page intentionally left blank

Part I: Research Desi

This page intentionally left blank

2

Thinking Tools

Anna Leander

In 1984, I moved to Paris to begin my undergraduate education at the *Institut d'Etudes Politiques de Paris* (Sciences Po). Sciences Po offered a series of seminars ostensibly to help foreigners (including me at that time) pass the entrance exam. What I remember from these is a chain smoking 'M. Thomas' doing his utmost to convey the message that Sciences Po was an elite institution, that entering it was like entering a 'gulag' and that only the best would 'survive' (his expressions). I also recall finding M. Thomas and his universe rather bizarre. A few years later, this was no longer true. I looked at French education in a new way just as Iver Neumann (in this book) looked at women differently after working with fur-coats. But more significantly, I had become intensely aware of the (often inarticulate) hierarchies and power relations of *practices*.

A year and a half after my first encounter with M. Thomas and the practiced hierarchy of French higher education, I came across the work of sociologist Pierre Bourdieu and more precisely his book *Distinction* (1984). (All references here refer to his work in translation but I recommend the originals in French, which tend to be considerably longer and more elaborate.) By that time, I was thoroughly puzzled by the idiosyncrasies of the hierarchies surrounding me as well as by the fact that those on the receiving end of these (students, including myself) kept accepting them. *Distinction* provided some clues, since it is an analysis of social hierarchy in France. But more significantly for a discussion of method, it contained a *vocabulary* for asking questions about power. These were embedded in a Social Theory of the grand kind: an updating of such classics as Marx, Durkheim, Weber, or Levi Strauss informed by philosophers such as Pascal, Kant, and Heidegger (such as Bourdieu 1996a, 2000a). No wonder I was impressed.

This chapter conveys some basic ideas regarding the ‘thinking tools’ this vocabulary provided that will be useful for applying any theoretical framework to empirical research. Bourdieu has attracted attention from all branches of the social sciences and the humanities, including international relations, resulting also in a momentous secondary literature. Clearly this is not an obscure method that seduced me because of my experience at Sciences Po. The chapter you are about to read cannot possibly ‘cover’ it or introduce an uncontested version of it. My presentation is selective, geared primarily towards the social science side and towards providing some practical advice based on my own experience in using it. Those who find Bourdieu’s particular tools potentially useful will also have a basis to find out more from his own work.

I will do this by discussing how the thinking tools relate to the key issues all researchers face when selecting and applying appropriate ‘methods.’ I begin with the kinds of questions that Bourdieu’s thinking tools are useful for raising and answering, namely questions about symbolic power and violence. I then discuss the conceptualization of the thinking tools in general terms, and proceed to highlight three crucial decisions to be made when ‘operationalizing’ these to answer a specific research question. The chapter concludes with a discussion of how to distinguish good research from bad (validity), with an emphasis on the centrality of reflexivity.

Research questions: ask about symbolic power/violence in practices

The method a study uses cannot be dissociated from its research questions. Methods serve a purpose. One does not drill holes with a hammer or fix nails with a drill. Similarly, when working in the social sciences it is important to acknowledge that methods can do different things. The method one chooses is related to what questions one is answering. Inversely, as anyone embarking on a research project (and any supervisor) knows, formulating a good research question is key to a successful research project. Methods textbooks explain that ‘good’ research questions are anchored in existing literatures and theoretical approaches. There is a two-way relationship between research questions, theoretical approaches and the methods tied to them.

Consequently, the first thing to consider about a method is whether or not it is useful for formulating and answering the kind of research question one wants to ask. The ‘thinking tools’ introduced here interested me precisely because they gave me a vocabulary for considering the

questions I found important about my Science Po experience, namely questions about symbolic power and violence in social practices. I have continued to find these significant in my work in international political economy and international relations, ranging from the politics of foreign direct investments in Turkey to security in Africa and in the West (see, for example, Leander 2001, 2002).

Asking questions about symbolic power amounts to looking at how 'symbols' (broadly defined) are an integral part of power relations (Bourdieu 1992). This seemed of essence in the 'gulag' that M. Thomas was introducing, but it kept striking me as an essential aspect of all power relations, including in the very hard material things. To stick with Sciences Po, there was clearly a strict hierarchy; there were dominated and dominating people. This hierarchy had some material manifestations (material rewards for success, written rules, sanctions, institutionalized humiliations) but the common understanding of education and of one's own role in the system seemed so much more important. It seemed to shape the material manifestations of power relations as much as (if not more than) the other way around. Asking questions about the working of symbolic power hence seemed an obvious priority. The thinking tools were helpful in that they directed my questioning towards three central aspects of these power relations.

The first of these was the extent to which 'symbolic violence' was an integral part of symbolic power. The power relations at Sciences Po could not have worked if the 'losers' of these relations had not themselves gone along and followed rules, which so obviously placed them at a disadvantage. As in so many other situations, the victims were their own perpetrators. Women perpetuate gender inequality, military establishments accept benchmarking practices favoring private security companies, development planners contribute to a displacement of the focus of development thinking towards security issues. Symbolic power relations rest on 'symbolic violence' where victims perpetrate their own powerlessness. Power therefore works all the more effectively as there is a degree of what Bourdieu would call 'misrecognition' or *illusio*, an idea with parallels in Gramscian and Foucauldian thought. For similar reasons, power is all the more effective when it rests on understandings which appear disinterested or unrelated to hierarchy, for example, based in science, culture, or art (Bourdieu 1993, 1996b). In my own work, technocratic competence, efficiency, humanitarian work, and local empowerment have been central for obfuscating power relations and symbolic violence. To 'discover' this, asking questions about symbolic violence has been crucial.

The second aspect of symbolic power relations that the thinking tools help focus attention on is the centrality of practices (what people do) rather than overarching discourses and representations (captured by what they say and write). At Sciences Po, rules were upheld more by what was *not* said and written anywhere than by what was. Power rested on the innumerable practices people engaged in without thinking much about it, just because it was the right thing to do, and they all somehow knew it. When you arrive as a foreigner, you notice simply because you do not know and (consequently) keep doing the wrong things. You would really like people to articulate the unwritten rules for you but if you ask, it turns out they cannot. For them, the rules are so obvious and natural that they do not seem to be rules but part of the natural world. Texts and discourses (and Sciences Po's written regulations) will of course reflect some of this, and you can capture this part by reading and acquiring a 'cultural competence' of the kind Neumann mentions (in this book). But the step from discourses to practice is a long one (see Dunn's discussion of the 'long conversation' in this book).

This brings attention to a third aspect of symbolic power highlighted by the thinking tools, namely its link to the material world (things like money, jobs, institutional positions, weapons, passports, or diplomas). Meaning and its practical implications change depending on the context. *What* you say matters less than *where* you speak from. The mystery of the minister is that her words can produce the material realities they purport to represent. But they do so *only* because of her position in social hierarchies. Similarly, the power of contemporary private security companies reflects not only the favorable bias towards them in risk and new public management discourses but also their links to policy makers, their evolving institutional role and their capacity to promote these economically. In addition, to some extent, what you say depends on where you speak from. As a student in Sciences Po, I did not count on having the same effect on our reality as our professors or as the minister of education. In fact, it did not even occur to me to try to have much influence at all. What I say (or not) is linked to my social position. This focus on material power and social hierarchies as an integral part of meaning production contrasts starkly with the 'internalist' focus of those discourse analysts who concentrate mainly or only on language. It has consequently been a key bone of contention between Bourdieusians and (some) post-structuralists (see contributions in Shusterman 1999).

To recapitulate, methodologies are linked to conceptualizations of the social world and so are the questions they are useful for answering. The approach introduced here is particularly helpful for asking questions

about symbolic power and violence in social practices. This may not sound terribly original. Discourse analysis, process tracing, and gender studies methods – just to mention some methods discussed in this book – claim to raise and answer similar research questions. However, as just underlined, I find Bourdieu's approach particularly helpful because of the specific focus it gives to these questions. It keeps questions about power in the center of the analysis. It directs attention to the centrality of the dominated in power relations. It is helpful for capturing the extent to which practices reflect and reproduce a mixture of economic, cultural, and symbolic power. With this specific focus comes a set of methodological tools. Consequently, the next steps are to get a hold of these in the general toolbox (conceptualization) and then decide how you would like to use the tools for your own purposes (operationalization).

Conceptualization: grab your thinking tools

A general conceptualization of the social world is an integral part of any methodology. It defines what to think about and what to look at (hence thinking tool). Methods rest on these assumptions about how the social world works. With vision come basic tools. Some authors in the social sciences become 'classics' because they challenge existing assumptions and make readers see the world differently. Luhmann, Braudel, and Foucault have made people think about how the social world works in novel ways. One cannot use Foucauldian discourse analysis or a Braudelean historical materialist analysis to answer questions about Luhmannian autopoietic systems. When Neumann (in this book) advises you to begin by carving out a 'discourse,' he has already equipped you with the basic thinking tool for analyzing the social world: not the carver but the discourse. Bourdieu-inspired methods rely on three such basic thinking tools: Fields, *Habitus*, and Practices (some would add *doxa* and capital). Indeed, earlier versions of this paper talked about the FIHP (Field, *Habitus*, Practices) method.

The first of these thinking tools is the field, the centrality of which leads some scholars to label the method 'field analysis.' In order to make sense of the social world, it is useful to think of it as divided into relatively autonomous sub-systems following their own logic. These subsystems are called fields but the general idea is rather widespread and reminiscent of Luhmann's relatively autonomous social systems. Sciences Po might be thought of as a field, relatively autonomous from the field of social sciences internationally, from the French economy, and so on. A field is defined by the fact that those who are in it share

an understanding (often unarticulated) of the rules of the game or the 'stakes at stake' in that given area of social activity. In that sense, the field is essential for understanding power relations. It defines what counts as advantages, or (social, economic, or cultural) 'capital' in that field. People's (or institutions') relative *position* in the social hierarchy in turn is defined by how much capital they accumulate. In diplomacy, humanitarian aid, banking, or Islamism, different forms of capital confer advantages. While central bankers may hold dominant positions in the field of banking, they may be subordinate in the humanitarian field.

Fields are only *relatively* autonomous. They exist in the context of other fields. This means that capital can be imported from one field to another. For example, Halliburton could import the economic and political capital it had accumulated in the field of US construction when it began competing for security contracts in Iraq. Of course, there is no guarantee that capital in one field has the same value in another field. Halliburton's political contacts to the Pentagon and the State Department were certainly more directly valuable than were its contacts to local administrators in Houston, Texas, when it moved into security contracting. There is an 'exchange rate' for capital. One might think of the struggle over its value in terms of the general struggle for power in society, and it is in this sense that Bourdieu uses 'the field of power.'

That fields are only relatively autonomous also means that the logic of a field is continuously shaped by the logic of other fields. Some fields are particularly important because they influence a great number of subfields; one might think of these as 'meta-fields.' Education, with its role in defining legitimate knowledge, is one example. The State, with its claim to a 'monopoly on legitimate symbolic violence,' is another. The shift in a meta-field sends ripples across a number of other fields. For example, the revalorization of neo-classical economics, including econometric modeling and degrees from the United States or Britain, triggered changes in most other fields, such as public administration, where these assets become valued and new public management thinking central. In turn, this shifts the positions and capital of actors in a range of subfields. In security, for example, private firms found themselves considerably advantaged. The meta-field of education has been crucial in reshaping the subfields of public administration and of security. These linkages between fields, and in particular the existence of meta-fields, are useful for understanding the broader (re-) production of power and domination in society.

Fields are not only static entities where actors occupy immutable positions according to their 'objectively' measurable capital endowments.

Fields are also dynamic terrains of struggle. People may seek to improve their own position by increasing their capital, they may strive to alter the field increasing the value of the capital they have or they may try to shift the boundaries of the field to alter both the value and the amount of capital they have. It is surprising that this struggle is not more intense and explicit. To explain this and to give substance to struggles that do take place, the second central thinking tool of the approach, the *habitus*, becomes pivotal.

The idea of the *habitus* is that while people have resources (capital) granting them a position from which to act, they also have taken-for-granted understandings, or 'dispositions,' that guide how they act. These are largely habitual and unreflected in nature, hence the term *habitus*. But they are essential for power relations. The *habitus* shapes strategies for accumulating capital and for reshaping fields or the failure to have such a strategy. But more than this, dispositions – such as eating habits, cultural interests, manners of speech, dress codes, and lifestyles – give shape to the body and body language. These become incorporated and embodied capital. Atatürk's dress codes (prohibition of the Fez and the veil, detailed dress codes directed at state officials) and the contemporary struggle over them are good illustrations of efforts to shift the value of incorporated capital and more profoundly of the dispositions going with them. Atatürk wanted a modern and Westernized Turkey. Present day Turkish Islamists wish a Muslim and independent one.

The *habitus*, like capital, is produced in specific fields. It reflects the values and discourses of a field, which in turn are shaped and reproduced by the people in that field. It provides the link between general structures and discourses – to which the Bourdieuan *doxa* is a rough equivalence – and the variety of practices they result in. Hence, the *doxa* is useful for the analysis of broad overarching understandings (such as Bourdieu's analyses of the state) or for the analysis of relatively undifferentiated societies (such as Bourdieu's analysis of Kabyl society in *The Logic of Practice* [1990]). However, to understand why a person or groups of people reflect general discourses in varied ways and why people follow the kind of 'strategies' they do, the *habitus* is a better tool.

The *habitus* is indeed an agent or group level thinking tool. As such it is subject to variation and change. A person is part of multiple fields in the course of their life. A person entering a new field (me entering Sciences Po, International Alert activists entering diplomatic circles) is bound to miss many unwritten rules and consequently appear clumsy and ill-adjusted. Over time, these rules become incorporated into the *habitus* of the person, whose behavior becomes less awkward. Alternatively, the

logic of the field might evolve so that the behavior is no longer at odds with its logic. Often both processes occur. Activists of major non-governmental organizations, such as International Alert, have learned the rules of international diplomacy and to some extent these rules are reflected in their *habitus*. At the same time, they have been major drivers of change in international politics. For example, their mere presence, which is at odds with traditional diplomatic state-based politics, has resulted in far-reaching changes in what actors can claim to be part of the field (extended to a range of non-state actors), what resources are valued (adding democratic resources, media power, and human rights credentials to military and economic might), and what understanding about international politics is taken for granted (such as in resolutions passed by the UN).

This takes us to the third thinking tool, practice. The basic idea with practices is that what people do rather than what they say is of essence. In part, this is so because a large share of their behavior is not consciously reflected but habitual and shaped by the position they act from. Practices capture the 'structuring' effects that shape action. (For a Foucauldian perspective on this issue, see Dunn and Gusterson in this book.) It is a way of capturing the reasons and situated rationality of action by replacing it in context. It is a guard against the very common tendency to impute a rationality to people (usually the rationality of the researcher) and then be forced to explain behavior that does not follow this rationality as stupid, irrational, or deviant, a tendency Bourdieu referred to as the 'genetic fallacy.'

More centrally, practices capture what people do in context, and this *relational* aspect of practices is of essence. We may be able to understand the action of International Alert in calling attention to small arms trade in the UN context by looking at its capital and the *habitus* of key members. However, we can only grasp the *habitus* and the capital if we think in relational terms. Moreover, if we want to understand the consequences of their actions for power relations in international politics, we need to place this action in relational context. We need to look at the practices of International Alert, how these are shaped, and, in turn, how they affect the practices of other actors in the field. Since practices are thought of as relational, they capture the overall pattern of interactions in a field and are differentiated from individual strategies of action.

This leads to a last essential point about practices: they are 'generative.' Practices create meanings, entities, and power relations. When International Alert enters international politics, practices are shifted.

It is not simply that power relations change because (given) people gain and lose in terms of some (predetermined) resources. Rather the resources and the people that count in international politics themselves are reshaped. A generative process is in motion. Similarly, contemporary political practices resting on opinion polls and media-mediated political action 'generate' politics as the aggregation of atomized individual interests on topics over which individuals have little to say and often few thoughts (see contributions in Wacquant 2005).

In practices, one can observe the relations of (symbolic) power and violence. It is hence not surprising that many consider 'practices' pivotal to the approach. They would argue that any Bourdieu-inspired study should depart from practices and build up an understanding of field and *habitus* from these. More generally, they would side with those who consider Bourdieu's work as key to the 'practice turn' in the social sciences. However, as pointed out above, the *habitus* and field have similar status for other scholars. My own position is that the three thinking tools are related to each other and work together. Perhaps this is because I first read *Distinction*, where the analysis is framed as [(*habitus*) (capital) \rightarrow field = practice]. But more likely it is because I have worked with all three thinking tools and find them all important.

To sum up, the toolbox of this approach contains three basic concepts for thinking about the social world: field, *habitus*, and practices. Using these thinking tools together is the basis for explaining and understanding symbolic power and violence. Many scholars consider one tool to occupy a more central and logically primary position. My own understanding is that they work together, that one can begin by using any tool. Moreover, most studies make more use of one tool than the others. Certainly Bourdieu's own work did; note the difference between *Distinction*, *Outline of a Theory of Practice* (1995) and *The Weight of the World* (1999). The decision of which tool to use and how much to take the two other tools out of the toolbox are decisions about how to employ the general thinking tools in one's own context. As this indicates, the third step, after asking questions and conceptualizing, is to operationalize.

Operationalization: decide on boundaries, level and scope

The thinking tools have been used to look at symbolic power and violence in practices ranging from those related to artistic production, the state, international law, elites in Brazil, the family, the suburbs of Paris, the media, European politics, and public administration (and elsewhere). As this diversity signals, there can be no firm guidelines for

what exactly to look at, what evidence to gather or in what kind of quantities (nor can I possibly list here all the fascinating secondary literature applying Bourdieu). Annual income, bonuses, thinking in terms of financial economics, interest in extreme sports and participation in professional meetings may be essential for understanding the field, *habitus*, and practices of investment banking but have little relevance for understanding those in the field of artistic production.

It is impossible to ‘operationalize’ field, *habitus*, and practices before the research. Fleshing them out in order to analyze symbolic power and violence is what the research is about; ‘operationalization’ is a key aspect of research. This said, if it is to work well, there are three central decisions to be made about the study: (i) where to draw the boundaries of the field; (ii) at which level to work with the *habitus*; and (iii) how to limit the scope of the study (possibly through a selective use of the thinking tools).

Drawing boundaries around the study to delimit the field and the practices at the center of the research is necessary: we obviously need to know what symbolic power/violence we are interested in. Yet, the stakes are high. The delimitation of the field both includes and excludes. The drawing of lines therefore shapes the analysis and its results profoundly. Consider two studies analyzing changes in international security after the Cold War. In one, the boundaries of the field are narrowly drawn around diplomatic practices (Pouliot 2003). In the other, the boundary is drawn to include the gamut of security professionals, including police, military, and commercial networks (Bigo 2005). The subsequent analyses differ in content, coverage, and style. And they reach opposite conclusions about the nature of change in international security. Pouliot argues that security greatly increased after the Cold War, as the bloc confrontation has been replaced by a security community. Bigo concludes that insecurity has greatly increased, as a consequence of the evolving practices of security professionals.

It is therefore important to be conscientious about the decision to draw boundaries. Mistakenly drawing lines may distract attention from essential practices and power relations, and hence obscure precisely the things the analysis purports to clarify. It is particularly important to watch out for two common pitfalls. The first is to draw the boundaries of the field so that the symbolic power/violence relations one aims at analyzing fall outside it. Although there is an international diplomatic practice and field, it may be a serious mistake to assume that symbolic power/violence in the definition of international security can

be analyzed in terms of it. The pivotal role of security professionals and their routine practices, for example, is entirely left out.

The second pitfall is to assume that links between a field and other fields deprive the field studied of its own logic. All fields exist in context. This does not make it impossible or meaningless to study them. The crux is to draw the line between the field and practices that are central and those shaping them from elsewhere. The practices of private security companies can be studied in terms of a field in its own right, even if this field is obviously tied to a number of other fields, notably fields of national security which shape the field of private security professionals and which these in turn influence. However, for the sake of a study it is of essence to set the boundaries of which relations of symbolic power/violence one wants to focus on.

This leads to a second crucial decision that has to be made: what level to work on, or more specifically, how to operationalize the *habitus*. At one extreme, one might work from the individual. Hence to capture symbolic power and violence in the Caucasus, Derlugian (2005) has constructed his research around the biography of Musa Shanib to clarify and explain the (sharply diverging) political trajectories of Checheno-Ingushetia, Kabardino-Balkaria, and Abkhazia. At the other extreme, one might imagine working at the level of the entire practice and field studied, as Ashley (1989) did in IR, where he argued that the shared (Realist) assumption, or *doxa*, that community in international anarchy is impossible resulted in a diplomatic practice blocking the possibility of 'global governance.' Both extremes have serious drawbacks.

Using the *habitus* at an overly general level makes the social world seem uncomfortably 'automatic and closed,' as Lahire (1999) rightly points out. It overemphasizes the structuring effects that weigh on actions. The variation in the *habitus* of different groups and people due to their social positions and past experiences is simply eliminated by fiat, as is the role of emotions in social relations, such as love, family, friendship, or enmity. If the *habitus* is merged with the *doxa*, it can no longer provide the link between general discourses, structures, and agency. Its role as a separate thinking tool disappears. Working with the *habitus* on the individual level is no more persuasive. Here the *habitus* becomes a collation of individual experiences and pasts, in which it is difficult to distinguish what is of more general utility for understanding the symbolic power and violence of social practices. The *habitus* runs the risk of being watered down to an individual history with limited analytical clout.

Consequently, I find the best strategy to be one of trying to work with the *habitus* at a level between these two extremes. More concretely, the *habitus* works best when differentiated according to key groups in a study, as in Bourdieu's analyses of the educational field in France (for example, *Homo Academicus* [2000b] or *The State Nobility* [1998]). This is also how it enters my analysis of the field of private security, where groups of individuals share a common general understanding of the stakes at stake but differ fundamentally in how this is articulated in their readings of the social world. This 'middle of the road strategy' can usefully be complemented by analysis at the individual level to retain the sense of depth in the study. Like the *Economist* uses boxes to detail an example, one can use examples to flesh out a point. I have often relied on extensive quotes from interviews with security contractors, job announcements, and advertisements by firms to make arguments about the *habitus* of contractor groups more tangible.

The third and final decision to be made is when to stop or how to limit the scope of the study. This is a central question for analysts using any method, and certainly in studies drawing on Bourdieu, it is an essential one. The empirically grounded theoretical set up easily produces overly ambitious studies. Evidence – including statistical data, biographical information, photographs, art, literature, classical texts, diplomatic archives, public speeches, newspaper clippings, and interviews (depending on the question) – tends to pile up but could always be completed with even more. This requires subjecting 'evidence' to a thorough analysis. Finally, writing and structuring the analysis is inspiring, but word limits, stylistic requirements, and the like quickly become a nuisance. This is one reason for Bourdieu's foundation of the journal *Actes de la recherches en sciences sociales*, where there were NO word limits and one could publish non-conventional material including pictures, art, and news clippings. It is probably also the reason *Distinction* is 660 pages and *The Weight of the World* is 1460 pages.

Most of us do not have the privilege of publishing books or writing dissertations of that length. Nor do many journals accept articles on the conditions of *Actes de la Recherche*. But even if we did, it is really hard work as Bourdieu often sneered at those who shun empirical studies. Hence my strong and articulate preference for good 'thick descriptions' (Geertz 1973) based on the analysis of a range of evidence is tempered by my self-preserving instincts and pragmatic approach to the needs of those completing their dissertations. I am persuaded that deciding on scope, as early as possible, is of essence.

I have tried both of the two most common ways of limiting the scope of my studies, and they both work fine. The first is to reduce the empirical focus of the analysis: focus on small *groups* of agents and practices. Restricting the scope of an empirical analysis does not have to be done at the expense of its theoretical ambitions. For example, in *The Social Structures of the Economy* (2005), Bourdieu uses an empirical analysis of the housing market in France to make a general theoretical point about the significance of social structures for the operation of an economy. The second way to limit scope is to work selectively with the 'thinking tools': instead of trying to provide an analysis based on field, *habitus*, and practices, rely on one of these, leaving the others in the background. This strategy is also used by Bourdieu in short lectures and essays, such as those in *Practical Reason* (2002), to concentrate on an argument. But perhaps the most important is to put strict deadlines and time limits. (Or as Gusterson notes, in this book, the grant money runs out.) That is a very unscientific but effective way of limiting scope, making sure that a project does not swell and become more ambitious than there is room for it to be.

The thinking tools introduced in this chapter are malleable. They can be used to raise questions and analyze power in almost any context. Yet, when using them in any specific context they have to be fixed. The field, *habitus*, and practices (*doxa* and capital) need to be given concrete and tangible meaning. This operationalization within a particular focus is a central part of the research process – no general blueprint can guide it.

Validity: work reflexively

As with all other methods, a Bourdieu-inspired approach needs to answer the basic question of how it distinguishes good research from bad. Since researchers using the thinking tools are left relatively free to apply these contextually, they will necessarily make different choices. How can one judge which account is better if two accounts, such as the studies of (in)security discussed above, reach different conclusions on the same question? But more centrally for most people, how can one assert the quality and validity of one's own work? The answer seen from the perspective of the thinking tools is simple: 'work reflexively.' Reflexivity hence becomes an integral part of the 'method,' which is consequently sometimes referred to as 'reflexive' sociology (Bourdieu 1985). I outline here three distinct understandings of what working 'reflexively' means for research practice and end with a note of how it is reflected in research

writing. (See Ackerly, in this book, for a complementary elaboration on these issues.)

At the most simple, working reflexively may mean reflecting on the quality and validity of the study in a methods textbook's sense. Evidence for a thinking tools study is similar to evidence used in any empirical work. It relies, variously, on statistical information, life span data, interviews, texts, photographic evidence, or pictures. Consequently, the usual standards apply. Issues such as the accuracy, adequacy, representativeness, and relevance of the information are essential for evaluating whether the 'evidence' of a study supports its conclusions. For example, if people are assigned positions in a field on the basis of information that can be shown to be false or irrelevant, that assignment is mistaken. If a scholar argues that an actor's position in the field of international security is greatly enhanced by the cultural capital linked to the mastery of Copenhagen School concepts and the educational capital that comes with a diploma from the Political Science Department of the University of Copenhagen, he or she is simply wrong. Similarly, a generalization about the *habitus* of private contractors based on the movie *Blood Diamonds* can be taken to task for generalizing on too thin a basis. Finally, the approach is set up to produce accounts about real-world symbolic violence and power and social practices. If these can be shown to follow very different patterns from those suggested in an account, it is wrong. These conventional checks on the validity of a study deserve being taken seriously (see the other chapters in this book for answers to these issues reflecting the authors' diverse thinking tools).

However, reflexivity at this level is insufficient. As all studies that take the role of meaning in social contexts seriously, studies made with the thinking tools approach have to answer some tricky questions regarding the status of the observer in relation to the observed. Specifically for this approach, it would be inconsistent to claim that the field of the social scientists was a field – the only one – where people did not have a *habitus*, did not struggle over positions, and were not engaged in practices producing symbolic power/violence. Since the approach makes no such claim, it needs a way of dealing with the tainting that the dynamics of the scientific field must give to its 'scientific' accounts of the social world (Bourdieu 2004).

This is where the second understanding of reflexivity comes in: working reflexively also means using 'epistemological prudence.' The basic idea is that researchers should 'objectify the objectifying subject,' that is, use the thinking tools to analyze themselves. This caution about the way knowledge is produced has direct implications for research. It is

the only road to limit the bias entailed in looking at the world from one's own perspective, such as me looking at the world of private contractors as a female French/Swedish Copenhagen Business School employee. It is also important in interacting with the people researched. The impact of my physical appearance, reactions, gestures, social status, and use of language tends to have an immediate impact on what interviewees say and leave out from their accounts. I cannot abolish this, just as I cannot, through reflexivity, eliminate my own bias in order to look at the world from nowhere. I can, however, do my best to limit its impact and also be aware of it when I analyze the results. This is 'epistemological prudence' in research practice.

Third, the researcher exists in a broader context, in a social world where privileged knowledge, such as that produced in universities, is of essence. Scientific practices 'loop,' to use Hacking's (1999) term, back into society and reshape its 'reality.' Categories and representations create their own social reality. Educational institutions are meta-fields that shape knowledge in other fields not only by producing categories but also by sanctioning careers. When scientific practices have looping effects, we need to be reflexive about what kind of 'reality' these research loops constitute. Epistemological prudence is a beginning. It can be used as a guard against the collective hypocrisy and self-delusion of assuming or pretending (rather than showing) that research agendas sanctioned by a scientific field are those most socially important. This is an obvious concern in the current context of the commercialization and internationalization of universities.

However, limiting the role of reflexivity to one of prudence is arguably both naive and irresponsible. Instead of 'prudence,' one needs reflexivity in a third sense: as a 'realpolitik of reason.' Purportedly neutral and objective scientific knowledge all too often presents unrealistic and unreasonable accounts of a world devoid of symbolic power and violence. However, precisely because knowledge is so central to the social world, these accounts play an essential role in perpetuating power by obscuring it. This delegitimizes work that effectively deals with issues of symbolic power. In this context, reflexivity (at least in Bourdieu's view) should be used to promote a realpolitik bolstering serious scientific work (with emancipatory potential) while denaturalizing, historizing, and unmasking (to use some clichéd expressions) the fantasy world of much of what counts as 'science.'

The first two kinds of reflexivity are relatively straightforward and palpable. They sit well with classical understandings of reflexivity, even if the notion of epistemological prudence gives it a twist. The realpolitik

take on reflexivity is more complicated. It runs against the idea of value neutral science with which most contemporary university education is imbued. It smacks of politicization. It has become (mistakenly I would argue) associated with Bourdieu's left-wing politics and hence understandably irritates people who do not share these. Ultimately, the question is one of alternatives. The alternative seems to ignore the looping effects of the sciences, unreflexively accepting their role. Any responsible thinking person (not only left-wingers) would presumably find this unsatisfactory.

By now, you are hopefully wondering how these three versions of reflexivity can possibly be stuffed into a research project. The short answer is that they cannot. If I write an article about intervention in Darfur, I cannot also include a full reflexive analysis of my own position in the academic field and the link of my study to the political context I am analyzing. There will most probably not even be much explicit reflexivity about the evidence used. There simply is not enough space; the reflexive grounding of the argument will most probably have to remain unarticulated. But then, that is the fate of most methodological and theoretical considerations that underpin a study of any kind. This does not diminish their importance any more than it does the utility of working reflexively, but it makes following the reflexivity of others more difficult. It also limits the time one sets aside to think reflexively. One may wish for a magical self-reflecting quill à la Neumann (in this book) to do the job, especially since most of us cannot spare the time to write the equivalent of Bourdieu's *Homo Academicus* to come to grips with their position in their own academic field or of his *Distinction* to come to grips with their position in society. However, I still contend that, even if the result remains unarticulated, working reflexively is sound advice.

Conclusion: thinking tools, dispositions, and irreverence

When I first read *Distinction*, I did not for a second imagine that I would one day be trying to distil some essential points about its 'method' into maximum 25 manuscript pages. The idea would have seemed absurd to me. For one, I did not picture myself as an academic. But more centrally, I did not think of it as a 'methods' book. I found the book interesting and helpful for strictly personal reasons but drew no link between it and my studies. As many students, I thought it essential to have neat and clear-cut concepts and methodological tools that simplify the world. The dense vocabulary, the shifting definitions, and the constant back

and forth between theory and empirical observation in Bourdieu's book definitely did not fit this understanding of a useful method.

It was not until quite a few years later (well into my PhD) that my frustration with the Procrustean beds of neat and clear concepts and methods that effectively stymied interesting research pushed me to draw on Bourdieu. By that time, I had come to appreciate the relatively open and malleable thinking tools. These did not work as the strict universal categories that I had once thought indispensable. They were integral to something more useful, namely a disposition for thinking about power and symbolic violence in context.

This chapter has communicated my bid for the substance of that 'sociological disposition' and more specifically my understanding of its methodological translation. I have insisted that I think it disposes analysts to raise questions about symbolic power/violence and, more generally, social hierarchies. I have suggested that thinking of the social world in terms of fields, *habitus*, and practices is integral to it. I have drawn on the work done by myself and others to point to some key decisions to be taken in the course of operationalizing these general thinking tools. And I have argued that it logically suggests the importance of working reflexively.

This distilling exercise is absolutely irreverent. I have imposed a strictly personal order, priority, and logic on a complex and multifaceted conceptual framework, which can of course be understood and used differently. Moreover, to satisfy editors and readers, I have eliminated much of the conceptual apparatus and ('all that French') vocabulary that expresses it in the process of simplifying. But then, Bourdieu was a great advocate of the irreverent use of theories – of 'writing with a theorist against that theorist' – so I may just be following the tradition I claim to write about. The bottom line is that if this makes what I have called the thinking tools more accessible, it will have been worth it.

Acknowledgment

Readers of the very different earlier versions of this chapter have made helpful comments: Stefano Guzzini, Jef Huysmans, Hans Krause Hansen, Karen Lund Petersen, Dan Nexon, Ted Hopf, Rens van Munster, Trine Villumsen, and especially Audie Klotz, her students, and my co-contributors to this book.

3

Feminist Methodological Reflection

Brooke Ackerly

Feminist inquiry is not reserved for women or even for those who identify themselves as feminists. It invites every scholar to revisit his or her epistemology and core conceptualizations throughout the research process. Feminist theory and methods provoke self-reflection, empower the researcher to explore new questions revealed by such reflection, and guide the research process in ways that are attentive to the power of knowledge. To assume that only feminists or women could do feminist inquiry would be to ignore the scholarship that feminist inquiry requires. If we assume that some people understand power through their struggles with power and not through their scholarly study of power, we belittle the scholarship of those who struggle.

Feminists share critical sympathies with post-structural, post-colonial, and critical scholars and with social movements, particularly women's movements, local and global. Among these, *feminists* do feminist inquiry particularly well because (and when) they are attentive to: (1) power in all of its visible and invisible forms, (2) boundaries and their potentials for exclusion, marginalization, and incomplete or superficial inclusion, (3) relationships of power and obligation (between people in different parts of the global economy, between men and women, parents and children, researchers and research subject, reader and audience), and (4) the role for self-reflexive humility in maintaining attentiveness to these concerns.

Distinguishing feminist methodologies from other methodologies is less important than asking how critical reflection can improve the conceptualization, epistemological assumptions, and research design choices required for any research project. Feminist methodology

encourages all scholars to acknowledge that there are hierarchies in our own scholarship and to acknowledge that our own inquiry is partial and ongoing. Without such consciousness, we are not only bad feminists, we are also bad scholars. In this sense, feminist methodology is not just applicable to questions about women (as Leander's contribution to this book illustrates). In addition to providing a huge range of particular tools for inquiry, the feminist contribution to methodology can be summarized as a tool for reflection that is guided by humility.

Whether a particular project calls for qualitative or quantitative methods, feminist inquiry entails reflection directed at all stages of the scholarly process. In coming up with a research question, we may ask whose interests are served by it. In conceptualizing our study, we may ask how language has historically conditioned the conceptualization of a problem. In operationalizing our variables and in collecting our data, we might use gender analysis. For example, an interview is a feminist interview when we reflect on the power dynamics between researcher and research subject in global context. (Gusterson's experiences with spouses, retold in this book, illustrates that men can do feminist interviews.) Publication is *feminist* when we attend to how our findings will be used and their effects on our research subjects.

Feminist theory has made feminist empirical work particularly challenging. Its commitments to exploring absence, silence, difference, and oppression have generated aspirations to do research that, if fully practiced, would leave many scholars forever in the field, always listening for new voices, always (respectfully) hearing cacophony, always suspicious of certain harmonies or recurring themes (Lazreq 2002; Dever 2004). Empiricists have to wrestle with this irony without allowing it to prevent them from doing their research. To allow such reflection to inhibit rather than inspire our research would be to perpetuate the invisibility of gendered absences, silences, differences, and oppressions, and the injustices that they conceal (Gluck and Patai 1991; Pettman 1992; Wolf 1992; Enloe 1993, 2000; Sylvester 1994; Wolf 1996; Stacey 1999; Staeheli and Nagar 2002).

In this chapter, I illustrate these concerns through my own work on human rights. I briefly introduce the critical ambitions inspired by feminist theory and then highlight the kinds of methodological challenges I faced when trying to be attentive to silence, marginalization, and absence. I reflect upon the relative merits of the options I considered and choices I made at important moments in the research process. In the third part, I offer a pedagogical and methodological tool, 'curb-cutting,' for inspiring one's own reflection.

Feminist inquiry in brief

Feminist inquiry is about revealing unquestioned differences and inequalities that conceal the exercise of power, including the power to conceal those differences and inequalities, and being attentive to the power exercised when researching these. All aspects of the research process are contestable. Attentiveness to the exercise of power extends not only to a field of study but also to its manifestation through academic inquiry.

Academic feminism has ethical responsibilities that reflect an ontological understanding of scholarship as *for* social change. Feminist theory educates feminist empiricists about the ethical importance of epistemological reflection at every stage of the research process. And feminist theory informs the methodological choices of any feminist researcher. Finally, most feminist research is self-consciously deliberate in its pedagogical purpose. The choices that feminists make about where and how to share and teach our scholarship are themselves methodological.

In empirical research, feminism encourages attentiveness to the challenges of seeing marginalization when the social, political, and economic authorities of a society render hierarchies either invisible or socially characterized as natural (Enloe 2004). It has been influenced by and has influenced many critical theoretical perspectives including post-structuralist, post-colonial, and critical theories (Fraser 1997; Narayan 1997; Ling 2002; Risman 2004) as well as movements for social justice (Collins [1990] 1991; Young 1990, 2001). At its best, feminist inquiry is attentive to the power of epistemological authority to mask political, economic, and social oppression as natural and accepted (Pateman 1988; Brown 1995; Ackerly 2000, 2008; Hirschmann 2003).

Such attentiveness to silence and inequality might well inhibit an empiricist from gathering or analyzing data by trapping her in a self-reflective mode. It might render a qualitative empiricist particularly incapacitated, afraid of exercising power over her research subject at every turn (D'Costa 2006; Jacoby 2006; compare Stern 2006). Juxtapose this potential incapacity with the emancipatory potential of feminisms (Agathangelou 2004). There is nothing emancipatory about fear – even the fear that one's own ideas may be corrupted by systems of power that one has internalized (Suu Kyi 1991; Ackerly forthcoming 2008). As Martha Nussbaum argues, 'In fear, one sees oneself or what one loves as seriously threatened' (Nussbaum 2001: 28). Threats may lead to emancipatory action, but threat itself is not emancipatory.

Moreover, while the feminist researcher may well have good theoretical and ethical reasons for leaving activism to activists (Bell and Butler 1999), there are no good ethical reasons for the feminist researcher knowingly to perpetuate the silence of the marginalized when she has the education and resources to reveal these (Smith 1999; Ackerly 2007b). And there are good ethical reasons for disrupting silence and marginalization. Yet, we know from critical reflection that feminists have also perpetuated some forms of marginalization, despite our best efforts. (Dunn, in this book, explores similar themes in the context of race and inequality.)

How then might the feminist empiricist proceed? Each choice she makes as a researcher should be evaluated as an exercise of power, just as it is understood to be an exercise of discernment. When the feminist empiricist attends to dilemmas that emerge during her research process, she may resolve them in the moment in order for the research to proceed. But they remain unresolved in the sense that at other stages in the research process, she may reflect back on earlier choices to note the epistemology that is privileged by those choices or the ways in which prior conceptualization has limited her ability to engage fully with the import of a particular dilemma. In the next section, I illustrate how I dealt with such dilemmas.

Methodological dilemmas in practice

From 1998 to 2001, I was a participant observer in online working groups of women's human rights activists (Ackerly 2001) and hosted with the Center for International Studies (CIS) at the University of Southern California a conference for scholars, activists, and donors. From these two projects emerged a puzzle: how could activists and other feminists who disagreed about so much understand their work for women's human rights as part of a shared project? Was this understanding an artifact of the resources (the online working groups funded by UNIFEM and the conference funded by CIS) that facilitated their dialogue? Was this the 'articulation' of a theoretical insight about human rights more generally? Or both?

Building on that work, I began a project exploring the notion of 'universal' human rights from a feminist perspective – a perspective attentive to absence, silence, difference, oppression, and marginalization that makes claims to universality both politically and theoretically suspect (Peterson 1990; Fraser 1999). Some of my methodological dilemmas affected the entire project, others only certain aspects of it.

Some could be attended to but never resolved; others could be provisionally resolved. I offer these not as representative of the dilemmas feminist qualitative methodologists face, but rather as opportunities for sharing how one feminist uses feminist methodology to think through them. As feminist researcher, I attended to:

- power in all of its visible and invisible forms,
- boundaries and the potentials for exclusion, marginalization, and inclusion to be incomplete or superficial,
- relationships of power and obligation (between people in different parts of the global economy, between men and women, parents and children, researchers and research subject, reader and audience), and
- the role for self-reflexive humility in maintaining attentiveness to these three concerns.

Jacqui True and I refer to these four as the ‘feminist research ethic,’ a shared tool of feminist empiricists. (We review the feminist work from which we derive this schema in Ackerly and True, forthcoming.)

For expository purposes, I discuss these dilemmas in an order that roughly reflects the chronology of a research plan. However, because of the way that the feminist research ethic guides our thinking about a given research dilemma, the feminist research process often requires deviating from the research plan and even retracing steps of the process. For example, dilemmas in sampling could make us rethink our question (see D’Costa 2006 for a particularly informative model of this). Each dilemma should provoke many questions and further reflection on related dilemmas. In order to share my work, I reconstruct key moments of methodological reflection, not to mask the non-linearity of the actual research process, but to enable others to comprehend the import of what are in retrospect decisive dilemmas in research.

My principle *question dilemma* was, should I ask the question, ‘Are there universal human rights?’ Is it a worthy theoretical enterprise for a feminist (who is attentive to the power of difference and the invisibilities of various exercises of power) to try to reason about *universal* human rights when most accounts of universals mask the particulars of privileged experience (MacKinnon 2006; compare D’Costa 2006)? In my preliminary work on the topic (in both the working group and the workshop), I came to know women human rights defenders. Many were working a third shift (after work and family care) to promote women’s human rights in their communities, many others were working at risk to themselves, and they seemed to understand

themselves as collaborating even while disagreeing. From the relative comfort of academe, I reasoned that I could start from the assumption that they were onto something. Therefore, the question was a means to interrogate power not to reify or conceal its exercise.

Having decided to ask the question of universal human rights, the *research design dilemma* was how to go about answering it. Feminist reflection about research design is similar to that of other fields at many levels. Feminists review the literature and consider the merits of various options relative to our research question. Like other researchers, we face decisions about locations, domain, sampling, data collection, data analysis, and publication.

However, some of the feminist considerations may seem unusual. For example, I considered broadening the use of the Internet as a source of insight, but I worried that continuing to focus on the online working groups would not yield an adequately diverse sample. Participants represented only a slice (and to a certain extent, disproportionately elite slice) of the women's human rights movement. In addition, the groups were becoming increasingly self-referential. Fewer new participants were contributing than when the groups were first launched. Generally, they were more privileged within this context: the relatively better resourced, those networked with the global North, those with local political networks, those with funding networks, those who were relatively more powerful within their organization (compare Ackerly 2007a). My question was not well addressed from the perspective of the privileged within their organizations and the privileged within the women's movement globally.

The ethical dimensions of research design may not always be anticipated by the researcher, but some can be. For example, I could have continued with the method of participant observation at meetings that I hosted. However, this approach raised ethical questions for the researcher-research subject relationship, some of which were related to the limited resources of women's movements globally (Clark *et al.* 2006; Ackerly 2007b). The meeting at CIS cost approximately \$20,000. Replicating that meeting would likely cost the same (or more, as there were many expenses associated with that meeting that were not incorporated into the budget of the workshop). Should some of the few resources for women's human rights activism in the world be diverted to my project? It was not clear to me that the under-funded work of women's human rights activists was best supported by their participating in meetings organized by me. However valuable they were to my work and to those

who were able to come, it was not obvious to me that this was the *best* use of funds.

I decided that I would do best to learn from activists in the women's movement transnationally and locally. Therefore, I looked specifically for activists marginalized in mainstream movements (for example, movements for street vendors) and for women marginalized within the transnational feminism movement (for example, women who were not part of the 30,000 women at the 1995 UN Conference on Women in Beijing).

Research design considerations include practical and ethical dilemmas around selecting research subjects, *the domain* of subjects from which we sample. Identifying my research subjects was an act of epistemological power that would have a definitive impact on my findings. What would be the best way to identify (to see and to locate) women in order to learn from them? What would be the domain of women activists? How would I sample among them in order to select activists to interview?

I had neither hope nor aspiration for a representative sample of the world's women or even of women's human rights activists. The movement is so huge, *geographically* diverse, and disbursed, and always shifting membership – all of the activists of the movement (the domain) would always remain unknown to me. To try to *get* a representative sample of an unknowable domain was a spurious endeavor. Moreover, my question could not be explored by a representative sample. Instead, I needed my inquiry to reveal different and competing ideas. Therefore, I needed to sample marginalized thinkers and those willing to voice disagreement.

Most systems that I thought of involved identifying visible organizations through Internet searches and my growing personal networks. These approaches would yield those who had made themselves visible to myself or others. For example, I might have studied recipients of small grants thereby relying on a third party to identify the organizations. Each of these approaches to identifying research subjects would have privileged the relatively powerful, even though some of those subjects might have identified themselves as relatively marginalized within global women's activism.

The international discussions leading up to the World Social Forum (WSF) and the Mumbai Resistance (which was a more anti-capitalist meeting that met across the street from the WSF conference) in 2004 indicated that these meetings might present an opportunity for me to witness people who felt marginalized in a range of contexts. By coming to the WSF, they would be exhibiting the willingness and ability to make

the effort to overcome that marginalization by planning to network. At WSF, I would be able to hear their thoughts about the possibility of networking when they were most aware of their similarities and differences with other groups because they were confronting them daily during WSF. I would also be able to be a participant observer of movement organizations' efforts to be more or less inclusive of other movements and organizations. Because of all of these opportunities to hear difference and thoughtful reflection on difference, I decided to interview participants in WSF and Mumbai Resistance.

Again, there was cause for reflection on privilege. Would not these be the elites of the transnational movement as well? Possibly. However, women were using WSF as a space to influence global progressive movements. Mumbai WSF was an opportunity for women to see if alliances with other progressive movements were possible (World Social Forum Panel 2004a,b,c). Further, though some participants in WSF 2004 stayed in comfortable hotels, the cost of participation (including travel and lodging) was feasible for many (a total of 960 rupees or less than 18 Euros). This meant that the meeting drew a large number of grassroots activists from India. WSF was a place where I could observe activists expressing their ideas and their agency. (I went to Porto Alegre the next year and sent graduate students to Kenya for WSF 2007.)

To be confident that my sample did not create a bias toward a shared universal view of human rights, I needed a *sampling mechanism* that would result in the study being informed by people who disagreed with one another or who were willing to voice disagreement with certain parts of the movement. A research assistant and I planned out which panels we would observe, seeking to identify a range of feminists. At those panels, we sought to identify women (and men) who could offer perspectives we had not heard before or who had asked a question that went unanswered in the panel discussion. Only four interviewees were selected because the person was known or referred to me. But this snowball sampling did not get any bigger than that because I wanted the anti-snowball sample: a sample of the people who were at some degree of critical distance from the snowball of transnational feminist activism. I needed to seek those whose opinions might differ from the main arguments heard by well-networked feminists. I was listening for cacophony.

What kind of *data* should I collect in order to be able to record these people's insights and bring them home for further reflection? I considered asking for life histories and other forms of open-ended interviews that would give the interview subject the greatest opportunity

to control the content she provided (compare Stern 2006; also see Gusterson's use of life histories in this book). However, these were activists who came to WSF to network, not to be research subjects. They were willing to take a break and reflect over a cup of tea or water, but they were not up for a long intermission in their own activist agenda in order to inform my academic agenda. With the ethical consideration of *their time* most central to my methodological choice, I decided on participant observation and an interview format, which treated the subject as the expert she was on the demands and possibilities for universal human rights from her perspective. I invited her to comment on some of the dominant threads of transnational feminism, to identify opportunities for networking and collaboration, and to offer accounts of obstacles.

Even while collecting the data, but most critically once home, I reflected on the ways in which *analyzing data* is an ethical practice: an exercise of power, of delimiting boundaries, of appreciating relationships, and an opportunity for self-reflection on all of these. If the interview subject was the author of her ideas, could I ethically do anything other than report them (Gluck and Patai 1991; Wolf 1992)? How was I to use her insights?

I decided not to take each interview as its own individualistic isolated 'text' but rather treat it as part of an ongoing dialogue with others, my other interview subjects, the others that the subject had had conversations with at WSF, the panels that she had participated in, even the interlocutors that she imagined. This was appropriate because participants were at WSF to dialogue with others. When they sat down with me, they were continuing a pattern that they were engaged in throughout their time at WSF – exchanging ideas with others. This dialogue was enriched when I went to WSF 2005 in Porto Alegre and to a meeting of feminists that preceded it. My understanding of WSF as a space for women's human rights activists changed as a result of these additional meetings, but the importance of reflecting on the differences among women's human rights activists was affirmed.

The dialogue was further enriched by sharing my insights. Though my interview subjects did not dialogue with each other, in my book I put their ideas in dialogue with one another. In a brief letter circulated to the interviewees, I shared a short account of the theory and the draft manuscript of the larger project. I solicited and received few comments from the participants. This stage of the inquiry was made particularly difficult by my choice of marginalized activist. Many of my interview subjects were difficult to find again. Their experiential reasons

for not communicating with me could not be known to me, unless they communicated them to me. One wrote from Afghanistan,

Thanks for your nice concern and yes!! You can publish my interview anywhere you want as you asked me in your last email.

As you know we are living in a society where religious extremists are capturing all the social and political Sector, and my activities are very democratic so while launching some of my projects and programs for women empowerment I am receiving different threats, and I have to be careful about my life because they are very good in shooting people like us....

(unedited personal communication, 25 February 2004)

In essence, she asked me to use her insights, whether or not I could find her. (Note that I have illustrated a common and yet unanticipated impact of analysis. It revealed the importance of dialogue and collaboration *for the analysis* of the data and therefore the need to have a second stage of data collection at WSF 2005 and another kind of data collected in the form of correspondence with interviewees.)

For most academics, *publishing* a piece of research is not a methodological decision; it is a professional decision about audience. In contrast, for the feminist it often is a methodological decision. In my case, what would be the value to the activists informing my inquiry of the theoretical treatise I was expecting to produce for an audience in my field, political theory? Taking their theoretical insights in order to produce what was for them a useless piece of scholarship, inaccessible in their familiar language, was an act of privilege inconsistent with the theoretical intent of the project. The epistemology and methodology of my project was supposed to break down the theory/practice dichotomy, yet the book produced for that project (and appropriate for my professional ambitions) could be inaccessible to the informants.

I dealt with this paradox of privilege in two ways. First, the book itself contains an argument unusual in political theory, a suggestion *to activists* about how they might change their work in light of the theoretical insights generated during *my* reflection on *their* insights. Second, *Universal Human Rights in a World of Difference* is not the final end of this scholarship. As I finish the book, I am beginning a research-scholar network whose central question is, what does it mean to do research *for* social justice (as opposed to *about* social justice)? We are developing a research practice in the area of global feminism to do

academic research that serves organizations or communities engaged in women's human rights activism broadly (and problematically) defined. This scholarship has led me into ethical relationships with scholars and activists such that in addition to my academic interests, my ethical commitments require that I continue related scholarship *and* facilitate the opportunities for others to do research that strengthens women's activism around the world (Ackerly 2007b).

The academic-activist relationship developed during this scholarship has also created a pedagogical obligation (again, an *obligation* on my part, not just an interest, though that too) to bring into the curriculum of my elite university in the US theoretical reflection and empirical data that shine light on the marginal spaces of global politics where so many of the world's women work so hard against oppression, unacknowledged, unrecognized, and 'invisibly.' This scholarship is intended to reveal that 'their' marginalization, oppression, invisibility, and absences are evidence of 'our' privilege and 'our' exercise of epistemological power. How do we teach ourselves and our students to do this work?

Curb cutting as a pedagogical tool

Because gender hierarchy so often gets institutionalized in ways that render it invisible, researchers need to comprehend marginalization through experience, not solely through sociological study. Of course, we cannot directly experience *others'* marginalization. So we need to teach ourselves to be attuned to the possibility that there may be experiences that are invisible to us. I describe this to my students as 'curb cut feminism.' This same pedagogical tool can be a stimulating tool for scholars' (not just women's) methodological reflection. In this section, I show how researchers can stimulate reflections on their own research process similar to those I described in the preceding section.

In my own classrooms at an elite North American university, where all facilities are technically accessible, I use the metaphor of curb cuts to reveal the challenge of making privilege visible. Curb cuts are the cuts in a sidewalk that turn a step into a slope and enable a wheelchair to cross a street or driveway. They are accommodations that make spaces designed to privilege access for certain people accessible to all. 'Curb cut feminists' cut ramps into the curbs of injustice – such as structural inequalities and the politics of misrecognition – and then we all, including immigrant minority women with dependent parents and grandchildren and trafficked camel jockeys, can live in a world that is more just, and promote justice for some in ways that enhance justice

for others. 'Curb cut feminism' also describes a theoretical perspective with a methodological commitment to be attentive to privilege and hierarchy not only as a subject of study but also as a *way of studying*. The metaphor may be best suited to North American university campuses where curb cuts, automatic doors, and blue disability symbols are part of daily life, but the aspiration to a form of emancipatory scholarship that works against *all* forms of oppression is important to many strands of feminist inquiry.

To illuminate the ways in which privilege gets treated as 'normal,' I ask my students to do basic daily activities – get coffee, pick up their mail, park for class – in a wheelchair. The purpose is not to simulate the perspective from a wheelchair (that cannot be done), but rather to make visible their own privilege that they did not see before. Through this activity, they gain a range of insights about (limits to) mobility created by the design of their institution and its accommodation systems. They notice steep hills, bumpy sidewalks, accessible side entrances and steps at the main entrances. One lesson that students quickly learn is that *accommodations* by no means mitigate the *privilege* of certain forms of mobility.

The exercise stimulates our desire and ability to notice additional forms of our own privilege, which we do not often notice. Further, by it I hope to cultivate the ability to reason in ways that make some of the experiences and challenges of those *unprivileged* by the basic spatial, political, economic, and social designs of our societies *visible*. Because I use the metaphor of curb cuts, one might misunderstand me to be privileging some forms of disability over others or privileging disability over other forms of marginalization. The point is not to identify a perspective to privilege, but to deploy a device that destabilizes the perspective from which we 'know.' Aware of the presence or absence of curb cuts, only one door being accessible, and of library books on inaccessible shelves, we are more attentive to the *privilege* of the norm. Attentive to certain mobility privilege, we start imagining improvements that would help *everyone*.

The ramifications of improved mobility are felt far beyond the initial focus on freedom of movement for those whose mobility is most challenged. We are able to see how redesign (and accommodation) can enhance the mobility of those pushing strollers and delivery dollies, those carrying lots of books, and those recovering from an injury. All are able to move more freely and safely. For example, parents with children in strollers can go to museums, which furthers their children's and their own education. However, curb cuts do not discriminate about the kinds

of mobility they enable. Access ramps may attract skateboarders or facilitate cyclists' going on the sidewalk to avoid automobile traffic. In this way, curb cuts might *inhibit* the mobility of everyone. Attentiveness to all forms of mobility *can* yield design solutions that enable safe mobility and better living for all, but it does not necessarily.

It is not enough *that* we pay attention to invisible privilege and marginalization; it matters *how* we pay attention to them. In designing our architectural landscape, we could rely on ourselves to do our best to think about the needs of people with disabilities, or to develop specific guidelines (like the International Building Code). We could design for those with mobility privileges and accommodate others or we could design so as not to privilege certain groups. No matter our approach, we cannot be confident that our designs (or accommodations) insure that our curb cuts worked for those in wheelchairs or that they made conditions better for all. Architects and city planners seek out 'rules' to help guide their choices, yet the insights of users are still the most valuable for noticing privilege.

Rather than relying on our best efforts – which can be fruitless or even harmful despite our intentions – we should develop *a method of thinking* that is always informed by the experience of people with mobility disabilities. Rather than relying on ourselves to do our best in thinking about the challenges and opportunities of marginalized people which could end up reflecting our own epistemological myopia (that is, our own unexamined view of what constitutes knowledge and of what data constitute evidence), we need *a method* for thinking about lived experience and for thinking about what lived experience tells us for theory and conceptualization, epistemological and empirical assumptions, research design and methods choices, and data analysis and publishing decisions (Ackerly 2000; Sandoval 2000).

Insights from feminist curb cutting inform my research and reveal the hierarchies between me and my research subjects and among my potential research subjects. Such reflections can help me carry out a research project that is self-reflectively attentive to power, boundaries, and relationships throughout the process. Applying these insights helps us see the political, social, and economic processes of normalization that sustain hierarchies. All sorts of values, practices, norms, and institutions impede, exclude, ignore, or marginalize some women, but not all women, nor only women. Curb cutting assumes that identifying and analyzing the conditions of oppression provides greater insights than those possible from positions of relative privilege. For example, I discussed above that I did not want to rely solely on Internet-accessible

research subjects or on women who were well integrated into transnational feminist networks.

Those who are 'most' or 'differently' oppressed may not be visible to the theorist or to the relatively powerful. Hence, for the follow-up interviews in 2005, I sought out women whom I perceived as silenced in the feminist meetings that preceded WSF 2005. Furthermore, the perspective not yet imagined is even more marginalized than the most marginalized perspective that one can imagine. I explored the silences in my own data: Who *was not* at WSF? What were my research subjects *not* telling me?

I put 'most' in quotes to indicate the need to reflect critically on the term. The point is not to privilege marginalization but to be self-conscious about the power exercised through marginalization and to be aware that the political claim to being 'most' marginalized can be used to challenge the patterns of hierarchy (Vélez-Ibañez 1983). As I describe in the preceding section, at every turn – from question through publication – the research needs to reflect on these power dynamics. Feminist curb cutting is a tool that any researcher can use to teach herself to be attentive to the politics of knowledge and to the power of privilege in her research practices. Attentiveness to these politics does not remove them, but it allows the researcher to unmask some and mitigate others, even as she enacts still others.

Conclusion

The methodological dilemmas I have discussed (and so many others) emerge as important in part because feminist theory says that the feminist researcher should worry about the ethical implications of the hierarchies of knowledge within which she works and to which she contributes. Attention to these requires not turning away from the discomfort associated with hierarchies of knowledge, but rather committing ourselves to looking for and to attending to them, always, already.

As I thought through these methodological dilemmas, I was working within and contributing to changing feminist theory. Feminists are appropriately worried about universalizing across differences. My work shows that we should be likewise worried about failing to work across differences because we are worried about universalizing. Feminism is hard work. Feminism cannot be a theoretical perspective that legitimates *not* attending to hard questions because the politics of epistemology are difficult to unmask. Feminist theory guides feminist empiricists to continue to seek out such dilemmas and to expose them.

Each research question is unique and each scholar will be faced with specific methodological dilemmas. I have offered my dilemmas and theoretically guided process of attending to and working within them as an outline in order to indicate the scope of feminist methodological reflection – from research question to publication. At no stage in the research process can the feminist, attentive to power, be confident that her research methodology has adequately interrogated the possibilities for absence, silence, difference, and oppression that the power of knowledge and research can conceal as well as reveal. However, that humility should not obfuscate her responsibility for the choices that she has made. Being ‘in the field’ brings ethical responsibilities. Taking that responsibility requires recognizing that ‘getting it just right’ is a privilege itself, one best shared. For many feminists, theory and empirical work is in the largest sense collaborative (Ackerly and True 2006). Such collaboration requires sharing our dilemmas, our imperfect efforts to work through them, and our partial insights.

4

Case Selection

Audie Klotz

For most researchers, case selection *defines* method: a few cases of a particular phenomenon make a study 'qualitative' but a lot of cases turns it into a 'quantitative' analysis. Usually a case is equated with a country, and there is often an implicit presumption that some sort of history will be traced. In International Relations (IR), qualitative method typically means a study of one or a few foreign policies, with a decision-making process to be traced at the micro-historical level (George and Bennett 2005). Yet for many questions, say, about globalization, countries are not necessarily the appropriate unit of analysis; economic systems might be. And historical evolution can happen at a higher level of aggregation, such as macro-historical changes in property rights.

Too often, the justification for a research design begins and ends with the rationale for the number of cases, obscuring key issues, such as the unit and level of analysis. In part, this is the result of the problematic conventional dichotomy between qualitative and quantitative methods. Would a project where the researcher uses statistical analysis *within* a single case study be qualitative or quantitative? From a comparative perspective, much of the work on American Politics looks just like that! Intensive analysis of a single case can employ all types of methodological tools without agreement on the degree of general, extensive, knowledge being sought.

Researchers need to remember that cases are cases *of* something. Well-crafted case selection takes into account the universe of possible cases and the logic of comparison implied by the research question. In this chapter, I will show how clarifying the overall purpose of the project and its theoretical framework broadens the rationales for single case studies, paired comparisons, and slightly larger studies. Often I will draw on guidelines by other researchers and suggest their publications for

further reading. Most of my illustrations will come from my dissertation-based book, *Norms in International Relations: The Struggle against Apartheid* (1995), because I can delineate explicitly the sorts of trade-offs and choices that rarely appear in published work.

Cases *of what?*

Appropriate case selection depends first and foremost on ontology, because any research question relies on core concepts. That brings us to the starting point for case selection: a case *of what?* As Leander (in this book) underscores, questions and concepts remain embedded in theoretical presuppositions. Quite often, these assumptions and subsequent propositions would benefit from clarification. Vagueness is not always the result of sensitivity to context and complexity! What are the key concepts that define a 'case' and what are the key dimensions that should be compared? These are not simply questions of finding indicators, although definitional decisions do hold implications for that stage of research design (Adcock and Collier 2001; Goertz 2006).

When I was formulating my dissertation, for example, I confronted the question of how to conceptualize 'race' in global politics. Was it an ideology? While otherwise quite useful, I found that this standard conceptualization underplayed contestation, and I was intrigued by the international controversy over South African apartheid, particularly the policies of neighboring states. Out went ideology. Alternatively, should I analyze the word race as a linguistic signifier? I found semiotics too focused on specific words, leaving out the material and social dimensions that ideology did capture. Was racism cultural? Yes, in a general sense, but the term 'culture' implied a dense immutability inappropriate for studying IR. I was looking for something less monolithic. Each of these formulations had advantages and disadvantages, but none seemed to capture how the global politics of race appeared to me in the late 1980s.

In the end, I opted to define race in terms of contending 'norms' of racial equality and racial superiority, situating my study in the context of regime theory. Responses to South African apartheid became *a puzzling case of international consensus* that challenged prevailing theories of cooperation based on rational calculations of material interests: Why would racial equality trump domestic jurisdiction? Adcock and Collier refer to this as the shift from a 'background' concept to a 'systematized' one (2001: 530–1; Goertz 2006: 27–57), moving the researcher from abstractions toward measures.

Yet conceptualization is not simply a one-way process, from general to specific. The way the researcher narrows a general concept in order to do empirical research also affects the formulation of the main question – not solely the choice of ‘indicators.’ Had I defined race through a different theoretical framework, my key question (and subsequent case selection) might also have shifted substantially. More influenced by Michel Foucault and Edward Said, for example, Roxanne Doty (1996) pursued similar questions about race through an analysis of hegemonic discourses, rather than regimes. She queried the constitutive role of race (in the imperial relationships of the United States with the Philippines and of Britain with Kenya), whereas I concentrated on a moment of contestation over it. We started with similar frustrations with the omission of race in IR theories, but our alternative theoretical frameworks led to different key questions, core conceptualizations, and subsequent cases.

Because IR theory lacked any standard conceptualization of ‘race’ at the time, Doty and I each independently devised a definition to put into practice. That potentially leads to the commonplace critique that case study researchers define concepts idiosyncratically. But the tendency to contextualize concepts need not be an insurmountable problem for comparing across cases. For instance, even an elusive concept like ‘regime’ has fuelled reams of insightful research on international cooperation and global governance, despite abiding definitional disputes. And applications of constructivism and critical theory have advanced in the two past decades to the point that researchers should be able to find enough common ground in definitions of race (though Doty might not agree with me on this). Most qualitative researchers remain comfortable with a moderately flexible set of characteristics, and many acknowledge the danger of ‘stretching’ a concept to the point that it loses its essence (Sartori 1970; Collier and Mahon 1993).

Still, case study researchers should avoid undue vagueness and would benefit from the series of questions that Goertz poses (2006: 30–5). He starts with a very basic question: What is the *opposite* of the concept? For example, democracy might be contrasted to authoritarianism *or* monarchy, depending on the research question. The opposite of racial superiority (of which apartheid was one manifestation) might be non-racialism (not accepting the existence of race as a way to categorize people) *or* multiracialism (not privileging one race over another). Another useful suggestion is to pay attention to the use of adjectives that modify key nouns, such as ‘parliamentary’ or ‘presidential’ democracies, or racial equality versus racial superiority. Since I did not

have Goertz to prod me before I got 'to the field,' I had to figure out some of these distinctions – and their *political* significance – as I parsed the sanctions debates.

How many *possible* cases?

After defining a case and the dimensions of comparison, the researcher needs to decide whether to analyze the full universe of cases or some subset. That requires identifying the 'universe' of cases. It may be small, such as 'world wars' or 'nuclear weapons laboratories' (Gusterson in this book) or large, such as wars or laboratories in general. Clarifying what would be a *non*-case helps for delineating possible cases. A non-case of a nuclear weapons laboratory could be a weapons laboratory or a civilian laboratory. A non-case of war might be violent conflict that remains within the territorial boundaries of one state, a militarized interstate dispute averted, or stable peace. It all depends on the formulation of the research question, although clarifying the universe of cases might, in turn, mean going back to reformulate the core research question.

In my efforts to understand how race affected policies toward South Africa, for instance, I had to decide whether to include all sanctions policies or select a few 'senders' (in the language of that literature). Initially, this task seemed straightforward: list the relevant international organizations and states, then decide the feasibility of including all of them in the analysis. I remain indebted to my dissertation committee for pointing out that only looking at those who adopted sanctions would have prematurely truncated the list of possible cases. My universe of cases quickly expanded, because I needed to include all the *debates* over sanctions to capture times when sanctions *might* have happened but were rejected. (Similarly, see Ackerly's discussion, in this book, of her difficulties trying to analyze marginalized and silenced discourses.)

The distinction between cases and non-cases may not be stark. And the gray zone may actually be more interesting (politically as well as theoretically) than the poles – Britain had mixed policies which critics did not consider sanctions at all – but we still need the full spectrum in order to identify its significance. In more formalized terminology, Mahoney and Goertz (2004) offer their 'possibility principle.' They rightly note that most non-cases are actually implausible and the subset of possible cases is much narrower. The trick is to figure out the difference. If you were lucky, like I was, you have a dissertation committee – or a colleague sufficiently informed but not vested in the outcome of the research – to keep you honest about plausibility. Those comfortable with the language of

variables will also benefit from reviewing Mahoney and Goertz's rules of inclusion and exclusion (2004: 657–8).

Some studies go even further to analyze the non-cases, such as policies that are *not* adopted (Price 1997; Tannenwald 2007). Indeed, depending on how the research question is framed, one study's case may be another study's non-case. The use of counter-factual scenarios further expands the range of potential non-cases, although their use raises a whole host of additional issues (Fearon 1991; Tetlock and Belkin 1996). Indeed, I do wonder why post-modernists are derided for challenging the objectivity of historical narratives yet made up counter-factual scenarios can be taken so seriously by social scientists. I prefer to treat these as theoretical formulations or predictions, rather than cases *per se*, because all researchers employ hypothetical 'what if' and 'why not' scenarios, either implicitly or explicitly. And that leads us to the next step in case selection, sorting out the logic of comparison employed in the project.

Which logic of comparison?

Delineating a universe of cases (including non-cases) does not tell a researcher *how* to analyze them, beyond some notion of comparison. Even a single case is not unique, otherwise there is no basis for calling it exceptional. However, the comparison might be against an ideal type. There are diverse ways that researchers parse evidence within a comparative logic. The type of question the study seeks to answer, in turn, depends on its underlying logic.

For example, King, Keohane, and Verba's (1994) controversial advice to increase the number of observations applies a statistical logic of theory testing: the larger the universe (or representative sample), the more persuasive the hypothesized claims about patterns between variables. Yet, for better or worse, redefining a key concept in order to create more observations may fundamentally alter the research question. World wars are not the same thing as militarized interstate disputes. If the research question really is about *worldwide* war, or *nuclear* weapons laboratories, the change may be unwarranted, and a study of only a few events or locations would be appropriate regardless of what the statisticians might think. For questions looking at wars or laboratories in general, though, expanding the study could be beneficial, since it reduces over-generalization from the experiences of great powers or scientists working on secret projects.

Making a choice between, say, five cases that lack the ideal controls versus a near-perfect quasi-experimental paired comparison that shifts

the main question depends on the researcher knowing the embedded logic of comparison. This *sounds* simple. Typically, social science projects seek to make somewhat general causal claims, while those drawing on post-modernism do not. Yet once we start delving into what exactly it means to make a causal claim, neither epistemological position proves to be so obvious.

There is no single 'scientific' logic. Some, inspired by physics, advocate deduction to generate hypotheses, usually followed by statistical testing. Others prefer focused comparison, because it mimics a test tube experiment in a chemistry laboratory. Biology and geology offer other templates. Policy-focused or intensive ethnographic research can also lead to general claims that might be reformulated as hypotheses to test and may convincingly disprove a prevailing theory in particular circumstances (see 'Single Case Studies' below; also Leander and Gusterson in this book). All researchers, therefore, would benefit from clarifying their analytical assumptions by asking themselves the following three general questions.

Does the study seek to test theories?

Avowed social scientists are not the only ones who put forth general claims; a study does not need to use the terms 'hypothesis' or 'variable' to offer theories that can be tested. Furthermore, theory-testing studies are only as good as their hypotheses. And as Ackerly underscores (in this book), there is no guarantee that factors, notably gender (and I would add race), have not been systematically omitted. Framing theoretical insights through the dominant scholarly discourse of testing propositions can lead to productive engagements, even if the initial studies do not use the vocabulary of variables.

For example, critical theorists have raised the visibility of 'omitted variables' such as gender and deserve credit for getting them into the 'equation.' See, for instance, Goertz's overview of how the addition of gender transformed the literature on the welfare state (2006: 88–93). In IR, Foucauldian notions of epistemic power have made in-roads in the past 20 years, as the limits of a materialist conception of power have become increasingly apparent. Not coincidentally, feminist approaches have gained legitimacy along with constructivism (Peterson *et al.* 2005). I adamantly refused to have variables in my dissertation, but these sorts of examples have mellowed me over the years, and therefore I encourage others not to react to scholarship based solely on differences in terminology (Klotz and Lynch 2006).

Are causal claims made in terms of conditions or mechanisms?

One of the most basic lessons about analytical inference is the differentiation between correlation and causation. Its application seems straightforward: use statistical analysis to identify patterns, and then select cases to illustrate which direction the causal effect runs (or figure out if another variable explains both). Case studies trace a causal process that links the proposed independent variable to the dependent one, in order to offer an explanation for why the pattern emerges. This research design has been the bread-and-butter model for mixing methods. More recently, the marriage of rational choice theory (which derives its hypotheses deductively, rather than observing statistical patterns) with historical narrative relies on the same process tracing approach to theory testing.

Elaborating on this as a social *scientific* basis for case studies, advocates of mechanisms have been contributing a lot to the burgeoning literature on qualitative methods (George and Bennett 2005; Gerring 2007). As with all approaches, a focus on mechanisms has its strengths and weaknesses (Checkel in this book). For example, such studies in IR usually demonstrate a chain of decisions by policy makers that presumably links the independent and dependent variables. But because any mechanism can be scaled up or down (Tilly 2001), no one can possibly test all the plausible alternatives. Once again, we are reminded that ontological assumptions about units and levels of analysis are critical.

Yet the problem runs deeper than researchers failing to test propositions about alternative processes. Causal chain narratives downplay contingency and contestation. Indeed, in his (otherwise quite useful) practical guidelines for historical research in IR, Trachtenberg (2006) *advises* writing historical narratives to emphasize the almost-inevitable nature of the outcomes, even if the preceding analysis does take into account alternative scenarios. But these chains of mechanisms, which focus on the presence or absence of factors, are not the only type of causal argument.

Probabilistic claims, articulated in terms of likelihoods, are based on conditions and the conjuncture of various factors at a particular point in time. This latter view shares a contingent quality that post-modernists favor. To see these distinctions in practice, note how Checkel and Dunn (in this book) employ documents and interviews in similar ways to different analytical ends. Checkel aims to create an historical narrative that positions social facts into a coherent story, whereas Dunn offers

a genealogy that highlights contestation over meanings. (Little 1991 offers a succinct overview of different types of causal arguments.)

Are constitutive claims adequately distinguished from causal ones?

My suggestion that probabilistic causal claims may have something in common with genealogy, through their similar emphasis on complex conjunctures, makes it especially important to understand the distinction between a 'constitutive' and a 'causal' argument. In post-modern critiques of social science, these two components often get conflated. Yet even Goertz, a quintessential social scientist, acknowledges the constitutive side of causal theories: 'concepts are theories about ontology' because they are about 'the fundamental constitutive elements of a phenomenon' which play a critical role in explanation (2006: 5).

One might think of this as treating the independent variable in a causal study as the dependent variable in a constitutive one. Foreign policies may 'constitute' identities, by inscribing definitions of Self and Other. Those identities, in turn, narrow the range of conceivable options. Identities thus play a 'causal' role in the sense of making certain choices more likely (and inconceivable ones, extremely unlikely). Formulating constitutive claims as conditional or probabilistic makes them causal *in a particular sense*. This challenges both post-modernists' claims not to be offering explanations and mechanism-oriented social scientists' claims to offer the only true proof of causal connections. Certainly constitutive claims *do not need* to be formulated as causal arguments. But I do think it helps to avoid tautology, which is particularly prevalent in arguments about the 'mutual constitution' of structures and agents (also see Hoffmann in this book).

There is at least one other advantage of *trying* to think of constitutive claims in causal terms: it encourages the researcher to think about what it means for a proposition to be 'wrong.' I do not mean 'falsification' in the narrow sense of largely discredited positivist standards for refuting theories. By 'wrong,' I mean that researchers benefit from thinking about what sorts of evidence might make a particular claim untenable. For example, the role of identity in foreign policy is a common theme in critical security studies. Quite likely, identity is defined in a way that precludes the possibility that there is *no* identity. ('What is the opposite of the concept?') But the question could be reformulated to *ask whether* specific interests conflict with specific identities, *rather than presuming* that the former derives from the latter.

Similarly, if the researcher assumes *multiple* identities, then the question might revolve around what a *dominant* identity might look like

(‘What adjectives modify the concept?’). Evidence might lead to the conclusion that a particular identity is *not* dominant; it might be subordinate (Neumann in this book), marginalized (Dunn in this book), or silenced (Ackerly in this book). Similarly, if the boundaries of identities are posited to be fluid rather than fixed, another common assumption, what would a relatively stable identity look like (Hermann in this book)?

My point is that researchers from many positivist and post-modern perspectives do parse evidence along similar lines, despite dissimilar philosophical moorings. This suggests potential for complementary insights, if researchers are willing to focus on their logics rather than labels. Recognizing this, we can put pluralism into practice in the selection of cases.

Single case studies

Often single case studies emerge out of an empirical puzzle. We see something that does not fit our expectations based on prevailing theories or conventional wisdom. A researcher, already knowledgeable about a part of the world or particular issue, may have some hunches about what is happening and perhaps some critiques of dominant frameworks. For example, my dissertation built on the observation that materialist theories were of little use for understanding why Zimbabwe risked so much to condemn racial segregation in South Africa. Simply put, apartheid should not have been an international issue if the Realist building blocks of IR, such as sovereignty and balance of power, were accurate. That observation, however, did not tell me what an alternative theory might be, nor did it tell me whether this one anomaly justified the wholesale rejection of Realism.

As my research question emerged, I readily found theoretical arguments that offered a plausible alternative – ideational – framework in Kratochwil and Ruggie’s (1986) critique of regime theory. However, this nascent constructivism did not offer a specific theory to test. Indeed, it resisted the whole endeavor of testing theories in the conventional sense! My research into why states and international organizations censured South Africa became a ‘plausibility probe’ to see if meta-theoretical arguments about the constitution of interests could be translated into empirical research. I presented Realism and Marxism as materialist foils to highlight key aspects of my alternative ideational framework, but the study itself was not designed as a *test* of any theory (evident to anyone who ventures back to read the dissertation’s theory chapter).

Plausibility probes are certainly not the only option. Single case studies *can* be used effectively to test theories if they fall into one of two categories. Some cases should be 'easy' for a theory to explain, yet it falls short. Others are 'unlikely' for a theory to explain, yet it does surprisingly well. There are various labels out there but these are the two general logics. Not all single case studies will fall into one of these two categories. Just the opposite: rarely will such a crucial case be available, but its analytical usefulness can outweigh many large-*n* studies.

An easy case can readily be confused with a plausibility probe, but the distinction is significant because each relies on the opposite logic. The exact same empirical evidence can contain more than one theoretical implication, but not all are of equal significance. For example, imagine that the evidence I gathered did show that norms could reasonably be interpreted as justifications for the pursuit of deeper material interests. An ideational approach would not be a better explanation of the censure of South Africa for racial discrimination – there would be little reason to reject Realism in favor of a new (and barely formulated) alternative theory. Yet a conclusion that Realism indeed could explain the putative weakness of norms would also not, in any strong sense, confirm the theory because South Africa was not a 'great power' (among other issues). Simply put, apartheid was a trivial case for testing Realism.

The second type of theory testing based on a single case is a 'least likely' scenario. Again, this should be distinguished from a plausibility probe, because the two may look similar. Unlike the easy case, a plausibility probe may follow the same logic as a hard case; the difference is the relationship between the theory and the empirical evidence. For example, if constructivist theory had been articulated in a less meta-theoretical way when I plunged into my dissertation, I might have framed it as a 'least likely' study because of the substantial amount of evidence in favor of materialist arguments (strategic resources in southern Africa, markets, and such). Other studies around the same time did directly target those theories by focusing on actors and arenas that prevailing theories considered most important: the World Bank (Finnemore 1996) rather than the Commonwealth, for instance (also the contributions in Katzenstein 1996). Not coincidentally, it was the part of my study on the *United States* that got published in *International Organization* (IO).

The value of single cases – perhaps more so than other selection rationales – depends in particular on the *status* of the theory that underpins it. In the late 1980s, constructivism had not been articulated to the point where, epistemological disputes aside, it could have been

tested – those ‘importing’ frameworks from other disciplines may face a similar situation. When I revised my dissertation for publication, I sought in the conclusion to translate my framework into more detailed claims that others could subsequently probe. In retrospect, I might have formulated these suggestions in terms of interests or identities as the ‘dependent variable,’ had I felt more comfortable with that vocabulary.

Now, if someone were to do a similar study, I would expect to see a research design that is built on the logic of easy or least likely cases, because the basic insight that norms or identities ‘matter’ is no longer novel. With the plausibility of the theoretical claim established, the value of doing additional single case studies (aside from the inherent value of knowing more about a particular place or issue) diminishes. As the circumstances that warrant the use of a ‘crucial’ case are limited, the research design questions shift to consider carefully paired comparisons or a larger set of cases instead.

Paired comparisons

Experimental logic makes carefully paired comparisons most acceptable for ‘positivists’ who aspire to test hypotheses. Yet given the infinite number of hypothetically possible variables across diverse levels of analysis, even carefully paired comparisons are inevitably easy to challenge. Outside the laboratory, as social scientists readily admit, ideal conditions will rarely exist. I refrain from saying ‘never,’ in recognition of a growing interest in field experiments, but these would be tough to apply widely in the IR context. Simulations, either with people or with computers, also offer potential insights, but they remain heuristics (Hoffmann in this book).

The closest approximation is the exploration of a single case over time, sometimes called ‘within case’ comparison, because it enables a researcher to hold many potential variables relatively constant. What might initially appear as one case turns into a comparative study. The best way to make this longitudinal approach work is when an ‘exogenous shock’ – the dramatic shift in an independent variable – enables the researcher to track closely what else does and does not change. Otherwise, there is nothing truly *paired* about breaking one case into component parts over an extended period of time. Simply ‘tracing’ the ‘history’ of a single case over time does not really take the logic of comparison seriously.

Yet even when there is evidence of such a sharp break, establishing historical stages remains difficult. Questions include how far back to

go and on what criteria to demarcate eras (see Neumann in this book). Quickly a comparison between two periods devolves into a longer study. And then researchers need to look beyond 'pairing' to tackle the difficulties of slightly larger small-*n* studies – the sorts of problems that George's notion of 'structured focused comparison' sought to alleviate (George 1979; George and Bennett 2005). I am not suggesting that we should abandon cross-temporal studies – we simply should not treat them as a special form of rigorous comparison.

The limitation which I find less frequently acknowledged in otherwise sensible discussions is whether paired cases are truly independent units or events. Tilly's (1984) notion of 'encompassing' comparisons comes close; some people use the phrase 'world time' to denote the importance of shared global historical context. Given increasing emphasis on 'globalization' across the social sciences, the question of inter-connections between cases (and not just variables) needs to be confronted explicitly. If countries, the most common unit of analysis, are not independent, then researchers need to figure out ways to control for external factors that may not appear as variables in the relevant literature. For instance, globalization has produced a new interest among comparative politics specialists in norms that diffuse to the local level, and they increasingly acknowledge significant cross-case interactions, such as emulation.

In this context, my South Africa study might be viewed as a study of the evolution of a particular norm (anti-racism). If there were a critical juncture, at which point one could claim that the norm emerged or consolidated, then a 'before' and 'after' study could be treated as a paired comparison, with most key variables either held constant or at least readily identified. In the United Nations debates over apartheid, 1963 marks such a turning point: the Security Council rejected a domestic jurisdiction defense in favor of a 'threat to peace and security' argument. However, in the Commonwealth, 1961 marks the key break: South Africa declared republican status and withdrew from the organization. We cannot do a structured, focused, *paired* comparison across these two organizations, because we cannot apply the same timeframe. Furthermore, the UN decisions took place in the context of prior Commonwealth debates, while those took place following earlier challenges to domestic jurisdiction by India going even farther back than the UN founding. We can learn a lot about, say, majority voting versus consensus by pairing these organizations, but they do not offer independent cases.

With few single cases passing muster as 'crucial' and so many inherent problems in paired comparisons, most qualitative studies fall into the

murky range of small-*n* studies. This gray-zone of 'more than two' but 'less than whatever is statistically significant' presents difficult terrain for case selection. The demands for detailed evidence garnered by a solo researcher are still possible but the results are inevitably more superficial. Some of these constraints can be alleviated through collaborative projects, but dissertation writers are less likely to gain funding for that, unless they work as part of a supervisor's larger project. Therefore, I assume that the trade-offs are faced by individuals.

More-than-Two but Not-a-Lot

Clearly, there is no single formula for dealing with multiple cases. Particularly for studies that start with an empirical puzzle, rather than a theory to test, some of the parameters of case selection are dictated by social realities and historical circumstances (see Dunn in this book). Yet I fear that research designs too often reflect the typical structure of a book: a magic trinity of three case study chapters, along with an introduction and conclusion, comprise a readable and reasonably priced volume. My goal in this section, therefore, is to get away from that trinity without tossing out the possibility that three case studies may indeed be appropriate.

For instance, in my study of international reactions to apartheid, I *could* have analyzed a wide range of international organizations and foreign policies. Yet it made sense to focus on the three communities in which South Africa had historically played a role: the international community (represented by the United Nations), the Commonwealth (initially as a Dominion within the British Empire but then during decolonization), and Africa (as a result primarily of geography rather than choice). Within each of these three communities, I analyzed the collective decisions of an international organization (UN, Commonwealth, and Organization of African Unity) and a key state within each group (the United States, Britain, and Zimbabwe). The result was, indeed, a reasonably priced book that I have been told is readable, as well as fairly convincing to specialists of each of these communities. I cannot complain too much about the magic trinity. But the choice of three communities was primarily inductive, the result of the historical legacies of South Africa's origins as a state in the international system.

Yet, on closer examination, counting the cases in my study is tough. These communities, taken together, comprise a single case of international cooperation (to condemn institutionalized racism in South Africa). In that sense, these are *not* six independent cases of sanctions

policies (in three instances across two types of actors). But if the theoretical focus were theories of decision-making, rather than regimes, they could be. In another sense, each pair of organizations and states could count as a case for tracing norm diffusion (but note the problem of interconnections discussed above). Going that route, the effects of sanctions on South Africa should also be considered a case, but it would not be a 'structured focused comparison,' because South Africa was the target while the other countries were senders of sanctions.

What is a confused researcher to do? My general advice, and my constant refrain in this chapter, is to remain mindful of the theoretical framework and core question, which do lead to reasonable conclusions about relevant cases at appropriate levels of analysis. And do not flee back to the world of single cases studies because they seem simpler; most of them lack analytical leverage. Yet that leaves my students profoundly unsatisfied, and I confess that I too remain uneasy. Other researchers (who also, notably, teach methods) have offered two directions for honing the selection of multiple cases: typologies and fuzzy sets. I am not yet convinced by either but both deserve serious attention from anyone trying to sort out this gray zone of Not-a-Lot of cases.

Typologies provide a fruitful path between the extremes of unattainable universal generalization and idiosyncratic contextualization. One of the advantages of a typology is that it offers an escape from the search for a crucial case or an elusive paired comparison by offering the possibility of comparing one or more cases against an ideal. Think about the adjectives often attached to concepts like democracy or war. These can easily be turned into descriptive or analytical typologies that differentiate forms of a phenomenon. And these typologies can be linked to constitutive or causal claims. One might explore a number of cases to illustrate the full range or concentrate on one cell, depending on the research question. (For elaboration and advocacy of 'typological theories,' see George and Bennett 2005: 233–53.) So far, so good.

Still, I advise caution, because it is seductively easy to draw up a two-by-two table for just about anything. That leads to a tendency to construct a dubious typology that justifies research that you already know you want to do or, especially for seasoned scholars, that relies heavily on research you have already done. I did just that for my dissertation: drawing on the sparse literature on pariah states, I identified two descriptive factors, which were only evident together in the South Africa case. I intended eventually to examine other historical examples (since

I claimed that none of the other contemporary cases were comparable). Although I dropped this convoluted foray into typologies for the book, I hesitate to dismiss the exercise completely. It prodded me to think more historically, which I generally was not inclined to do. And my favorite comparison, the Confederacy during the American civil war, inspired me to write a spin-off article comparing the abolitionists with the anti-apartheid activists (Klotz 2002). In the end, I still think that *insightful* typologies can help us avoid some of the difficulties of comparisons.

Another option gaining followers especially among those seeking to bridge quantitative and qualitative studies is Ragin's notion of 'fuzzy sets,' also known as 'qualitative comparative analysis' (Ragin 2000). Since his approach is full of technical terminology, I will simply mention here a few of the overarching goals that might encourage skeptics to take an initial look at some of his guidelines.

Rather than force complex concepts into rigid conceptual boxes, the notion of fuzzy sets accepts some inherent ambiguity. Concepts comprise a cluster of key characteristics, but no single feature is essential. Thus a 'case' of something includes some, but not necessarily all, of these core dimensions. Here is where the logic of Boolean algebra comes into play, and along with it, specialized terminology and formal notation. Anyone familiar with on-line library searches knows that typing 'and' gets a smaller number of hits than 'or' – it is the same logic. One might link this to typologies, for instance, by defining ideal types in terms of the most exclusive features ('and') while recognizing that cases will evince a subset ('or') of those characteristics (Goertz 2006: 84).

I find this logic appealing, because it helps me wrestle with a basic empirical question: should South Africa be considered a democracy? By the standards of the late 1700s, it certainly should – show me any political system based on universal suffrage at that time! By the standards of the late 1900s, its parliamentary elections without adequate representation clearly did not satisfy most definitions. Since electoral dynamics among white votes did play a significant role in the transition to inclusive democracy, the existence of certain features of democracy should not be overlooked. Also, there would not be much to the democratic peace literature if we used universal suffrage as a necessary feature for defining democracies in earlier times. Fuzzy sets move researchers away from essentialist terminology – which is also a major goal of constructivism and critical theory. Whether it can deliver on this potential for building conceptual bridges without getting mired in the jargon of its formal notation remains to be seen.

Conclusion

Perhaps because I grew up in a family of chemists, I never doubted the value of case studies. Paired comparisons come closest to controlled test tube experiments where one chemical agent (potentially) alters a reaction. No correlation will offer anything as compelling in terms of *causal* inference. Yet, I also never had illusions about the practice of science. For me, the laboratory was not the idealized space that philosophers contemplate; it emitted a distinctive aroma and was populated by human beings. And sometimes scientists – just like any other humans – have been extraordinarily successful in propagating ideas that subsequently appear quite ridiculous (Klotz 1986). I have never expected social reality to mimic molecules, because people are not objects. I also appreciate that scientists, like ethnographers, find some of their greatest insights while looking for something else.

Qualitative researchers of my generation had little to offer in terms of a methodological rationale. Scholars oriented toward theory testing easily dismissed our single case studies as ‘thick description,’ caricaturing Geertz’s famous 1973 essay of that name. I distinctly recall an awkward job interview situation that followed along these lines. Fortunately, we have come a long way in the past 20 years. The significance of single case studies for theory testing is still debated, but it is better understood. And its significance for theory building is widely accepted. Greater attention can now be paid to the messy middle of more than two but less than whatever is statistically relevant. Let me reiterate that these are research design questions that barely begin to address subsequent methodological questions of how to do the actual empirical study within the cases. The remaining chapters in this book do that exceptionally well.

Acknowledgment

Wagaki Mwangi and Deepa Prakash provided insightful written feedback on a short turnaround, for which I am especially grateful.

Part II: Classic Quali

This page intentionally left blank

5

Discourse Analysis

Iver B. Neumann

As a 22-year-old, I held a stray job fetching bundles of fur from minks, seals, polar foxes, blue foxes, and red foxes to show the buyers at Oslo Fur Auctions. The work altered the way I saw the world in other realms. Most striking was how one of my main interests at the time, looking at women, acquired a new dimension. Where I had previously focused on shape and movement, attire now became an important factor. My new interest in fur coats changed the way I sifted what I saw. Psychological experiments confirm my personal experience. For instance, children shown a cup with the handle turned away none the less drew a cup with a handle, because cups, by definition, are supposed to have handles. That is, the children drew the model, not simply what they 'saw' as a result of light waves hitting their optical nerves. People sort and combine sensory impressions of the world through categories (or models or principles). Language, as a social system with its own relational logic, produces reality for humans by mediating these sense data.

These examples highlight that perception is mediated by aesthetics, sexuality, morals, or other modes (Bauman 1992). In order not to forget that these meanings are socially reproduced, discourse analysts call them *representations* – literally re-presented. (I will concentrate on the preconditions for and jobs undertaken by representations; see Dunn's chapter in this book for more detailed discussion of analyzing the components of representations.) Representations that are put forward time and again become a set of statements and practices through which certain language becomes institutionalized and 'normalized' over time. They may be differently marked in terms of how influential they are. In the United States during the Cold War, 'dove' and 'hawk' representations of the Soviet Union were both institutionalized, but so was the (even less changing) representation put forward by the American Communist

Party. When people who mouth the same representations organize, they make up a position in the discourse. Like representations, positions may be dominant or marginalized in various degrees.

Demonstrating institutionalized discourse can often simply be done by proving that metaphors regularly appear in the same texts. In my study of European discourse on Russia, for example, I found a representation, which stressed that Russian females had been raped by Mongol and Tatar males for centuries, and that this had fostered a particularly wild and barbarous people ('scratch the Russian, and the Tatar will emerge'). This representation began to form fairly early, reached a peak in the inter-war period, and then lived a very submerged existence in European discourse. In the Baltic states, however, it was very strong indeed throughout the Soviet period and into the 1990s. The more such things may be specified empirically, the better the analysis. The ideal is to include as many representations and their variations as possible, and to specify where they are to be found in as high a degree as possible.

The first research task is to show the affinities and differences between representations in order to demonstrate whether they belong to the same discourse. But repetition does not preclude variation or gradual re-presentation, so discourse analysis also seeks to capture the inevitable cultural changes in representations of reality. For example, in the late 1980s, Russia was obviously heading for challenging times, and I reckoned that this would entail wide-ranging changes in relations with Europe. My basic idea was that, regardless of period, Russia's relationship with Europe had not been straightforward, yet it seemed set to remain central to Russian foreign policy as well as to Russian self-understanding. I wanted to be able to say something general about prerequisites for Soviet/Russian foreign policy in a situation where so many things seemed to be in flux.

Discourse analysis is eminently useful for such analysis, because it says something about why state Y was considered an enemy in state X, how war emerged as a political option, and how other options were shunted aside. Because a discourse maintains a degree of regularity in social relations, it produces preconditions for action. It constrains how the stuff that the world consists of is ordered, and so how people categorize and think about the world. It constrains what is thought of at all, what is thought of as possible, and what is thought of as the 'natural thing' to do in a given situation. But discourse cannot determine action completely. There will always be more than one possible outcome. Discourse analysis aims at specifying the bandwidth of possible outcomes. This works the other way, too; discourse analysis may also start with a specific outcome

and demonstrate the preconditions for it happening, demonstrating concurrently that the outcome might have been different.

To map these patterns in representations, discourse analysts examine utterances. They may be texts (written statements that do some kind of work in a context). However, any sign – a semaphore, a painting or a grimace – may be analyzed *as* text, because it conveys meaning in a particular context. Since we ‘read’ societal processes as the functional equivalent of texts, one may, for example, cull data from ethnography (see Gusterson in this book). Due to limits on length, I will focus in this chapter on written sources. (For an example of discourse analysis of ethnographic data, see Neumann 2007.)

Acquiring a certain cultural competence is a prerequisite for discourse analysis, as for most qualitative methods. After discussing the need for basic language skills and historical knowledge, I divide my lessons for method into three concrete steps. First, one needs to delimit the discourse to a wide but manageable range of sources and timeframes. From these texts, the analyst then identifies the representations that comprise the discourse, taking into account censorship and other practices that shape the availability of text. Finally, to explore change, one uncovers layering within the discourse. The more actions that the analysis may account for by demonstrating its preconditions, and the more specifically this may be done, the better the discourse analysis.

Prerequisite: cultural competence

I always encourage students to draw on extant knowledge when they choose their topics; it saves time, and they start out with a competitive advantage. It also provides a degree of ‘cultural competence.’ For example, I had done my conscription at the Norwegian Army Language School, where I studied Russian. Then I lived in Russia for half a year and did university courses in its history and foreign policy. All this gave me a certain cultural competence when I set out to research Soviet discourse as a doctoral student at Oxford (later published as Neumann 1996). I knew the Russian language, genres of relevant texts, and something about the general social and political setting (such as when Russia was at war with other states that it considered to be European and the extent to which European history and language were taught in schools).

This cultural competence enabled me to use tools of discourse analysis to demonstrate variations in meanings and representations. The more in-depth the general knowledge, the easier the specific research. For example, I knew that many Russian newspaper articles were divided into

two parts: a first part that repeated the so-called main line, then a part that dealt with new material that still had not been sorted in relation to and assimilated by that dominant representation. What is crucial for the discourse analyst is the separation of these two parts by *one* codeword, *odnako*, which is best translated as 'however.' If one knows such conventions, the reading of texts becomes easier: I could rush through part one, which is a simple re-presentation of an already known reality, and concentrate on part two. Similarly, the expression *en principe* in French signals that one is putting forth a representation which one generally shares, but from which one nonetheless is going to deviate.

Of course, some things may be learnt on the job. As a British-trained Norwegian Russian specialist, I needed to work at mastering phrases like 'to go' and 'drag it through the garden' to buy a hamburger in the United States. But there are other things that you have to know before you can start. When I turned to the analysis of discourse in the United States, it was inconceivable for me not to know references such as 'I have a dream' (a speech by Martin Luther King, Jr), 'beam me up, Scotty' (a line from the television show *Star Trek*) or 'I pledge allegiance' (to the flag). The point is that a researcher needs a basic level of cultural competence to recognize the shared understandings that create a common frame of reference, which makes it possible for people to act in relation to one another.

Let us not forget that the analyses we write up are written *for* somebody. What is adequate cultural competence for a specific discourse analysis hangs, among other things, on whether the resulting analysis may tell the intended readers something new. Ideally, a scientific text should tell every conceivable reader something new. That is a situation that is very rarely reached, however. The world is full of researchers who produce texts that do adequate jobs in adequate settings because they are new *in those settings*, and not necessarily anywhere else.

There is a trade-off with cultural competence. Culture appears to be shared. Close up, it turns out not to be. Phrases may mean a number of different things, or they may be used without the user knowing all their cultural references or implications. The challenge is not to get naturalized – not to 'become' part of the universe studied – but to denaturalize. If you are a native speaker and know a culture as only a native can, then you do not have that marginal gaze where things look strange enough to present themselves as puzzles. You will also lose touch with your own biases. You become what anthropologists call 'home blind.' For example, I once submitted an analysis of US foreign policy discourse which used a quote from the then chairman of the

Senate's Foreign Relations Committee, a key Republican senator of long standing, as representative of American discourse. The two reviewers, who wrote flawless American English, objected to my treatment of him as an elder on the grounds that they considered him out of touch and a crackpot, respectively! These readers were definitely 'home blind.' It is fully possible to do discourse analysis in the culture you know best, but you still need some kind of distance. You cannot be *too much* at home.

An historian or anthropologist would at this point ask, whose representations, whose culture? (See also Leander and Ackerly in this book.) We are talking about cultural competence regarding the culture that spawns the representations to be analyzed, not necessarily for other related cultures. When I was done with my discourse analysis of Russian representations of Europe, I noted that I had documented what I held to be so much arcane and downright silly Russian representations of Europe that I felt I owed it to the Russians to analyze European representations of Russia as well, presuming that just as much arcane and silly stuff would crop up. (It did; see Neumann 1999.) For that analysis, I needed neither Russian nor much knowledge of Russia. Instead, it was important to know German, French, and English. It was a problem that I could only cover Spanish and Portuguese representations in translations. But I still felt warranted in talking about European representations of Russia, for there were strong regularities between German, British, French, and Scandinavian representations of Russia at any one given time during the last 500 years that presumably could be generalized to 'Iberian representations.'

As in any other research, this lacuna has to be stated, and it will serve as a challenge to new researchers. (I have tried, so far unsuccessfully, to get a doctoral student to write about Iberian representations of Russia.) Methodologically, this points to the importance of being explicit about your sweep: the broader it is, the more general knowledge you need, and the less risky it is to leave lacunae. But great care should be taken here. No good Russianist would assume cultural competence about Serbia, and old cultural competence from the Soviet era may not necessarily be applied to Ukraine after its formal political separation from Russia. Knowing the ever-changing limits of your cultural competence may be as important as knowing its contents.

Step one: delimiting texts

Discourse analysts read texts. But what texts? In certain cases this is a simple question to answer. If one is to study party systems, then party

programs, election laws, and articles as well as speeches by party leaders are typical primary materials. Still, the quantity of material is usually enormous, especially if one includes the secondary literature. It is crucial to draw some lines, but problems of delimitation are inevitable. The choices applied to each individual discourse analysis always have to be defended. For example, if one repeatedly finds statements such as 'scrape a Russian and the Tatar will appear,' it would be mistaken to omit representations of Tatars from an analysis of Russian identity.

A given discourse cannot be entirely detached from all other discourses. They are ordered and scaled in relation to one another. Russian identity, therefore, must be studied as something Russian and something non-Russian. However, which relation or relations to study – between Russia and Asia, Russia and Europe, Russia and Germany, Russia and Tatarstan, Russia and the Jews, Russia and the feminine – is not given. Ideally, all should be covered. In practice, that is rarely possible. The choice of which relation(s) to single out may be theory driven (let us see what happens if we bring a feminist standpoint perspective to the study of Russian identity and look at the constitutive role of gender), utilitarian (I need to illuminate the identity aspects involved before I can get a handle on Russian-German energy relations; how do Russians think of Germany in general?) or ludic (my own favorite: why is it that Russians treat me the way they do? This must have something to do with general Russian ideas about Europeans.).

Insofar as politics is a struggle between named groups and people, politics is conflict. Conflict should therefore attract the analyst of political discourse. One will often find direct references to texts that are being attacked. It is usually apparent who is attacking whom. When there is such a racket, it is because something new is happening, something that is meeting various attempts at limitation from those who dominate the discourse (see Lukes 1974).

However, the pursuit of commotion can be a methodological problem, since realities are maintained by the frequent repetition and confirmation of representations. The absence of commotion does not mean that the discourse in question is non-conflictual. One has to use more time and mental energy to work out how and why things remain unaltered. Concentrating on the texts that produce the greatest racket might mean that one automatically privileges the dominant representation, which usually will be the loudest (Wæver 1999). Some texts remain unpublished when censorship is successful. Challengers may remain undetected for other reasons, including socially distributed lack of writing skills. One may also turn this around: publications that only

repeat or incrementally expand the main representation tend to pass relatively quietly. If one fails to detect these processes of power, then the analysis easily becomes a shallow one of the boundaries of the discourse and its domination.

Also, social and political life is full of cases where somebody writes something new and intriguing, with no immediate reception whatsoever. It may simply be that the text is so new or different in relation to what already exists that it goes unnoticed for this very reason. There are existing texts as well as future texts that will suffer this fate. If a text from a relatively obscure source becomes central – as did Francis Fukuyama's 'The End of History?' in *The National Interest* – then it is a research task to demonstrate how the text overcame the odds.

Some texts will show up as crossroads or anchor points, such as short government treatises outlining policy (called white papers in most English-speaking countries). These are called canonical texts or monuments (compare Laclau and Mouffe 1985 on nodal points). In my dissertation research, I was actually able to identify the textual canon by starting with the secondary literature, because it proved to be well informed. I took the 'monuments' to be the works that were generally cited in the secondary literature. I read these works, and indeed I found that they tended to refer to one another. This, as well as the negative finding that there were few additional central texts, confirmed them as monuments.

It is useful to select texts around these monuments, since monuments also contain references to other texts, which again pointed me to others that were related. One discovers that some texts are 'canonical,' in the sense that they have a broad reception and are often cited. If one identifies these texts, reads them, and then reads the central texts that these texts in turn refer to, soon one is able to identify the main positions and versions. In most contemporary Western nationality discourses, for example, the representation of history for political purposes is widespread.

However, it is not always possible to go back to antediluvian events, so one must delimit the timeframe. For example, once I had my dissertation topic, I read up on the secondary literature in order to identify cut-off points. An obvious one would have been the coming to power of the great Europeanizer Peter the Great in 1694. In order to trace discourse in more depth, I chose the Napoleonic wars that really brought Russia into the heart of European great power politics, and treated the period from 1694 to 1815 cursorily as a prehistory. The other cut-off point presented

itself during the work, as the Soviet Union split up in the autumn of 1991.

In specifying the sweep of the analysis, it is also important to keep in mind your reader. I later did a discourse analysis in my native Norwegian on Norwegian representations of Europe (Neumann 2001). I tried to tackle the question of home blindness by going way back in time – who is really ‘at home’ in the Middle Ages? In this case, the main intended reader was an informed Norwegian. I therefore needed to be fairly detailed in drawing up representations from the last 50 years. Yet I did not present context that was already fairly well known, which would not be particularly interesting to the prospective reader. When I did a shorter version in English (Neumann 2002), the intended readers were different, so I dropped detail and filled in context. A doctoral student in Europe, who has little idea who his readers will be, will tend to write differently from an American student, who has a committee from the outset. And how do you weigh writing for your examiners against writing for a more general audience that may also be interested in the texts? There are authorial decisions to be made – different strokes for different folks; broader ones for non-specialist foreigners, dense professors and academics working in outer disciplines.

Participants themselves also delimit their discourses. For example, medical diagnosis relies upon the definition of diseases and syndromes, upon which doctors draw. Analyzing the struggles over these definitions, and the process of getting them registered as such, form part of the research. If the chosen discourse is *international* intervention (to distinguish from medical interventions), then the struggle over the characterization of certain policies as ‘humanitarian’ is decisive. The main task is to dig out the production of specialized knowledge. In analyzing Norwegian human rights law, for example, there will be a number of relevant texts in legal journals and government policy papers. One can compare related professional discourses in other countries. However, the connection to general political discourse may not be explicit.

Some texts can acquire importance from the medium through which they are published. For example, a private letter from the 1830s threatened the dominant Russian representation of Europe after it got a wide reception through the circulation of copies in the saloons of St Petersburg, even though the author was declared mentally ill and incarcerated. It is important to bear in mind the values which different media give texts. If one is to carry out a discourse analysis of peace operations in the 1990s, it is important to distinguish between those journals that aim at operative milieus (*Foreign Affairs* or *Survival*), those that are

written for a more general audience (*International Affairs* in Europe), and those that are mainly read by academics (*International Peacekeeping*).

But what if there is a Russian letter or unpublished manuscript from the 1930s, unseen by more than a handful of people, which projects a representation of Europe that makes my analysis incomplete? In terms of the history of ideas this would be very interesting, precisely because of its originality and its lack of reception. Its discovery would provide a more accurate definition of the borderline between possible thought and the communication of that possibility. In terms of politics in the 1930s, however, it would be a non-event, because the analysis concerns texts that are socially communicated.

What if it turns out that there are a number of texts that are systematically overlooked, which jointly document that there was a main representation that previously had not been included in the analysis? In the area of women and war, one can at least imagine the possibility that a systematic reading of all available sources on the national service in Norway written by women would result in a revision of previous views of the national service institution (see Ackerly's chapter on subaltern discourses in this book). Mikhail Bakhtin's concept of genre is useful.

Genre carries its own memory, in the sense that every text relies on its predecessors and carries with it their echoes. If previous analysts have for some reason overlooked an entire genre, then it is an important research task to cast light upon how this has happened. This will change the way we remember a given historical sequence and is politically relevant to today's situation. Excavate one text on women and war, and you have an idiosyncratic voice and an indication that a group has not met the preconditions for action to make itself heard. Excavate many, and you have documentation that an entire group has been silenced. It is also possible that there are too few texts published, making it difficult to get started. One can carry out a discourse analysis of material that has not been in general circulation (for example, of classified material). If the reason for the lack of text is the novelty of the specific discourse, with, for example, only newspaper articles existing, it is possible to include a small text-based analysis of this material in an analysis that also draws upon other methods of data collection, for example interviews, surveys or participant observation.

When does one have enough material? The ideal situation is that one covers a maximum of eventualities, by reading as much as possible from as many genres as possible. Foucault insisted that one should 'read everything, study everything.' This is not feasible in practice, and there will therefore always be a risk that some relevant texts are not included.

However, almost regardless of the extent of the discourse, relatively few texts will constitute the main points of reference. Therefore at some point one has to be able to decide that one has read enough, even if one has not read everything. *Only* if a text emerges that cannot be subsumed under one of the main positions must the analysis be adjusted – or perhaps even rewritten entirely (see Hansen 1997).

Step two: mapping representations

A discourse usually contains a dominating representation of reality and one or more alternative representations. Discourse analysis therefore is particularly well suited for studying situations where power is maintained by aid of culture and challenged only to a limited degree, that is, what Gramscians call 'hegemony.' Structuralists and post-structuralists disagree over whether one can take a small part of the discourse and read it as symptomatic of *all* representations. Post-structuralists find the notion of a latent structure simply too deterministic. One must think flow, not control.

The task is to search out and identify these various representations and possible asymmetries between them. The analyst accepts and works with the inherent conflict between representations. Monuments frequently position themselves in the discourse by referring (adversarially or sympathetically) to texts that were previously considered monuments. Reading monuments in Russian foreign policy discourse, for example, helped me identify adversarial representations (for instance, 'Europe is vital, we should learn from it' versus 'Europe is rotten, we should isolate ourselves from it'), since these texts, often written at the same time, referred directly to one another. The advantage of a marginal position emerges clearly here for setting up an inventory of representations.

Researchers question how uncertain or challengeable a given representation is. The limits of discourse are inscribed with varying means and degrees of violence. If there is only one representation, the discourse is closed. This of course does not mean that it is not political, because it takes a lot of discursive work to maintain a situation where this representation cannot be challenged openly. If moves to do something new by the text-writer are not successful, it is not necessarily because the discourse is successfully policed. On the other end of the spectrum, the field can be said to be open if there are two or more representations and none of them are dominating. (See Leander's chapter in this book on the boundaries of Bourdieuan fields here; historically Bourdieu formulated his theory among other things as an extension and correction of

Foucauldian discourse analysis.) Yet it is difficult to imagine a discourse that is entirely open or closed over time. Social relations will always be in some degree of flux.

There is a second problem in addition to specifying the discourse's degree of openness. On the one hand, the number of permutations of relevant signs is endless, so the range of meanings is in principle infinite. On the other hand, politics involves contestation between relatively clearly defined positions, which compete to find resonance among a number of carriers. Thus it is desirable to identify these positions. Typically, one position will be dominant, and one or two other positions will challenge it on certain points. The dominant position will either present itself as being the way things have 'always' been (for instance, a democrat: humans are born free) or hark back to an idealized beginning (a democrat: Athenian democracy broke out of benighted despotism). Terms mean different things in different epochs, but carriers of a position will tend to tap the advantages of having a long (and presumably dignified) history by acting as if this were not the case (Koselleck 1988).

It is important that the discourse analyst start with the representations themselves – the stories of how things have 'always' been like this or that. For example, Athenian democracy was hardly a democracy by the lights of the 21st century. Neither was the ante-bellum United States. Arguing that every man is born free and has rights while having a number of living beings around who visibly are not born free and have rights (as slaves, or women, or children) reveals that the discourse is not open to the possibility that 'man' may be someone other than an adult, white male. Within the boundaries of his own political discourse, thus, it was not a problem that George Washington remained a slave owner throughout his adult life.

However, a good discourse analyst should also be able to demonstrate that where the carriers of a position see continuity, there is almost always change. Because of the nature of politics as a structured activity between groups, a discourse is politicized precisely through the evolution of two or a few patterns of meaning, which is the discourse analyst's task to uncover. It is possible to distinguish between the basic traits of such a meaning pattern (what unites the position) and varieties of it (what differentiate it).

In principle, the discourse will carry with it the 'memory' of its own genesis. Showing how each text is made possible by the preceding texts, often it is possible to find a prehistory to the main representation. It is, for example, hard to think of Stalin's funeral oratory for Lenin without having the model of the Russian Orthodox oratory in mind.

Methodologically, this is significant because, as a given representation establishes itself in the discourse, one should go back to find 'pioneer texts' that foreshadow it. This allows us to make a prediction: if a new main representation of Europe surfaces in Russian discourse during the next years, more likely than not it will be churned out of material that is already present in the discourse.

There are a number of formal and informal practices that determine which representations are allowed into the discourse, and that make it possible for the analyst to map meanings. Among the most obvious are legal systems and censorship, whereby sanctions against violating the boundaries of the discourse are threatened explicitly. An example: in Norwegian nationalist discourse of the 1990s just using the word 'race' activated a set of sanctions, foremost among which are laws that prohibit what Americans call 'hate speech.' The fact that there is no comparable Norwegian concept for the phenomenon, and that the American term is used regularly, are data for a discourse analysis of 'race.' (See also Klotz's discussion in this book of the concept of race in case selection.)

One can also examine what kind of self-censorship different types of mass media apply and what deviation it takes to provoke more formal sanctions. Legal verdicts on the borderline between incitement to violence and freedom of speech, and the debates surrounding it, would be one of several clues. To study nationalist discourse in the Soviet Union in the 1930s, where every newspaper, radio, and television station sifted what was printed and broadcast, one must start by examining the formal censorship instructions. Thereafter one might look at what unpublished and imported texts circulated, and what incidents resulted in Gulag sentences.

One should not overlook cultural artifacts with a widespread, so-called popular culture (see Dunn's chapter in this book). Discourse analysis is, for example, a useful way to examine film, understood as text. Rather than looking at museums, one can look at the reality production that happens in soap operas. If one is to examine the reality of 'Germany' in British discourse, then in addition to cases such as bilateral political discourse, EU discourse, and so on, it will also be of interest to look at representations of Germany in magazines, pulp fiction, and imported B movies (where it is still not unusual to find narratives where German Nazis are the crooks).

I would argue that the discipline of International Relations is not at present paying enough heed to artifacts of popular culture, but such an analysis must be situated, in the sense that one must be able to point out the inter-relation between representations of, say, Germany

in popular culture and political discourse about Germany. How does popular culture appear in and relate to political discourse? To what degree do representations from the former result in truth claims in the latter? 'Situating' (showing where something can be found, where it is *in situ*, 'in place') can be specified as proving inter-textuality between expressions, texts, and discourses (see Neumann and Nexon 2006).

Ethnography and discourse analysis are similar in that they pay, or should pay, a lot of attention to how the analyst is situated in relation to the data. In the 1980s, a key development in ethnography was an intensified attention to the writing up of the ethnography, and this turn was directly inspired by discourse analysis (Clifford and Marcus 1986). Typically, however, discourse analysis would splice data collection methods such as fieldwork or memory work with the analysis of written texts. It would also typically turn to written texts first, and think of other data collection methods such as interviewing as complementary or substitutionary.

Certain analysts are more formal in their mapping than others. I see heuristic value in being stylized. When discourse analyses are highly formalized, however, I always ask myself whether the reason is a need to appear social science like in order to get published, or whether it is actually an urge growing out of the text itself, whether it is necessary, and whether it is a market-driven or a scholarly necessity.

Step three: layering discourses

Not all representations are equally lasting. They differ in historical depth, in variation, and in degree of dominance/marginalization in the discourse. The third task for the discourse analyst is to demonstrate this. The production of gender is an example. There are a number of biological and social traits (diacritics) that line the boundary between the sexes, from the presence of ovaries to ways of brushing hair away from one's eyes. Few can be counted as unchangeable. However, some are more difficult to alter than others. It is easier to neutralize the gender-specific aspect of the sign 'unremunerated domestic labor' than 'childbirth.'

At this stage, some discourse analysts would cry foul, because they would like to insist that everything is fluid, and that nothing should be reified in the analysis. I agree that everything is fluid in principle, but the point here is that not everything is *equally* fluid. Furthermore, it is impossible to analyze something without reifying something else. Indeed, as my initial example of the child perceiving the cup is meant to bring out, it is impossible to see and to live without reifying things.

We have to subsume new phenomena into already existing categories in order to get on with our lives. Arguing that everything is equally fluid makes it impossible to analyze something in its social context. It also goes against what seems to be the very physiological preconditions of our existence as *Homo sapiens*.

Certain representations in a discourse will thus be slower to change than others. Signs that are 'good to think with' (Lévi-Strauss 1963) and representations of material objects will often be among these. However, now physical reality turns up. Put in everyday speech: material objects are difficult (though not impossible) to 'explain away.' But for the study of human behavior, this is not a problem. As Laclau and Mouffe illustrate,

An earthquake or the falling of a brick is an event that certainly exists, in the sense that it occurs here and now, independently of my will. But whether their specificity as objects is constructed in terms of 'natural phenomena' or 'expressions of the wrath of God' depends upon the structuring of a discursive field

(Laclau and Mouffe 1985: 108).

Meaning and materiality must be studied together. It is possible to take as one's starting point for a reading of a social event, such as the reasons why Sweden went to war in 1630, that there are a number of material 'facts,' including archaeological objects. Any valid representation of the social event must relate to and at the same time study the various representations of the social event without having to hunt some kind of 'truth' about it *beyond* accounting for these objects (see Ringmar 1996; Neumann 1997). The question is what the scope or degree of social construction is in the relationship between 'fact' and 'representation.' We should expect greater 'inertia' in the representation of material objects than that of other things, but this still does not ensure the place of the objects in the discourse.

This issue also lays bare the metaphors on which the discourse approach rests. Foucault wrote about archaeology and genealogy, the basic idea being that of things emerging, with some things remaining the same, and others changing. An archaeological site will contain certain artifacts that tell of continuity – there will be shards of pottery and traces of funeral rites – and these will vary with the period. But, in a particular site, certain things will remain stable whereas others change. The key, in archaeology as in social analysis, is to specify what changes and what does not, and how. The same is true of genealogy, the basic meaning of

which is that you start with one human and trace his or her ancestry. You will tend to find people who become less and less interrelated to one another the further back you go. At some stage, all they have in common is that they are all the ancestor of that particular human.

If some traits unify and some differentiate, it is reasonable to think that the traits that unite are more difficult to change. For example, if one chooses to study German identity, one will find endless variations on which things are thought to be German. If one looks at the question of how the state is related to the nation, the range of meaning will be lesser, perhaps only covering two possibilities: one, that the nation defines the state by being its cultural carrier, *Kulturnation*, or second, that nation and state are both anchored in citizenship, *Verfassungspatriotismus* (see Wæver 1999).

In my doctoral thesis, I approached this question of layering by postulating explicit and implicit family resemblances across time. The element of Europe as a place to learn from was in evidence at all points in time since the latter half of the 17th century, except for the High Stalinist period (two decades from the early 1930s onwards). In later work (Neumann 2004: 21), I formalized this step by drawing up a model of Russian discourse on Europe across time, using three layers: basic concepts (state, people, and so on), general policy orientation (isolation, confrontation, learning, and such), and concrete historical examples (pan-slavism, Bolshevism, early Yeltsin years, among others). At the level of the broad historical sweep, such a mapping of preconditions for action is the endpoint of discourse analysis. As should be clear by now from the discussions above, however, there remains endless work of specification on different constitutive relations, close-ups of specific time periods, tailor-making of the analysis to illuminate specific (types of) action, and so forth.

Conclusion: a discourse analysis toolkit

If one should fashion such a thing as a discourse analysis toolkit, it would perhaps look like this. Tool one would be a carver that would carve texts out of the social world. Tool two would be an equalizer that makes other phenomena (for example, a semaphore, an ad, a body) into material to be analyzed on a par with texts. Tool three would be something like a herding dog that would group these phenomena together based on them being about the same thing. Tool four would be a slicer, cutting the phenomena into different representations of the same thing. Tool five would be some kind of optic device that would

make visible the meaning dimension of the material phenomenon to its users. It would come with a grading spectre that could demonstrate how easy it would be to change the different layers of a given phenomenon. And finally, the only one that I would really like to see on my desk, tool six would be a self-reflecting quill that accounted for my own weighting of the phenomena of which I wrote as I wrote.

The point of such a tool kit would be to help us understand how the seemingly unchanging and 'natural' stuff of which our social worlds actually emerged as a creation of human history. Discourse analysis makes the social world more transparent by demonstrating how its elements interact. By demonstrating that things were not always the way they appear now, discourse analysis makes us aware that they are most probably changing as we speak. In order to account for global politics, therefore, it is not enough to study what one clerk wrote to another, how statesmen pontificate about the policies they pursue, or the technological changes that make for different kinds of warfare. The study of the meaning which these different phenomena have to those concerned also has to be included, and this means that discourses should be accessed at many different points.

Representations are constitutive in determining what is sensed and communicated, but they do not necessarily come with 100 percent built-in guides for action. If one has, for example, mapped Russian discourse on Europe, one has demonstrated several preconditions for foreign political action, but one has not necessarily cast any light directly upon the specific processes in the determination of such action (see Neumann 1996). A representation can make room for several different actions, and its carriers can be more or less conscious in their relationship with this representation. An analysis of representations of Europe will thereby not constitute an exhaustive analysis of Russian foreign policy. To do that, one must not only systematize the analysis of those sanctions that follow deviance, as I have already mentioned, but must also look at a number of other aspects of the materiality of discourse.

To the extent that a fuller understanding of where we are and how we landed here is helpful in getting us somewhere else, discourse analysis may be 'useful' for solving problems. But it is not your first choice in a tightly scripted situation, such as answering why state X went to war against state Y at point Z in time. Rational choice may be fine for that, even though the assumptions of the two approaches are very different indeed. An analyst may use discourse analysis in order to study how structures produce agents, and then decide to 'freeze' agents at a specific point in time, for example at the outbreak of war. The analyst may then

change tack and analyze the outbreak of war drawing on social choice theory. Such splicing of methods is highly unusual, among other things because few analysts are fluent in such diverse methods, but also since the analyst's own identity may be so tied up to one particular method as to make the very thought of mixing methods appear as sleeping with the enemy. To make a self-reflective point, why this is so may be studied by drawing on discourse analysis.

Acknowledgment

I thank Karen Allen, Stephanie Smith, my co-contributors and especially Audie Klotz and Deepa Prakash for a range of positive and productive comments on earlier drafts.

6

Historical Representations

Kevin C. Dunn

I am most interested in how certain social identities are constructed, and how they make certain practices possible but others unthinkable. Like Roxanne Doty, I examine 'how meanings are produced and attached to various social subjects and objects, thus constituting particular interpretive dispositions that create certain possibilities and preclude others' (1996: 4). I am less interested in 'what' questions, since these often prompt historical narratives that mistakenly assume a simple linearity of events. I am also less interested in 'why' questions, which tend to assume that a certain set of choices and answers pre-exist. Rather, we should investigate how those options and the larger possibilities of action get established. Doing so allows for greater understanding of the processes and interactions within international relations.

Choosing to explore these questions raises another fundamental 'how' question: How does one actually investigate structures of knowledge, such as social identities? How does one collect and analyze appropriate data? Because humans make sense of the world by navigating the social understandings that make reality knowable, researchers must employ interpretative methods. In doing my research on historical representations, I focus in various ways on language, ideas, and culture, particularly as they contribute to the creation of structures of knowledge during specific historical moments.

In this chapter, I discuss the various theoretical and methodological issues I encountered while researching my dissertation on representations of Congolese identity, which was later published as *Imagining the Congo* (2003). In the first section, I explain what I mean by historical representations, why it is important to study them, how they are linked to broader discourses, and why a deep historical analysis is needed. Employing a contextualized 'thick' description is useful for gathering

and analyzing data, I argue, but not without limitations. The rest of the essay is dedicated to a frank discussion of how one does this type of research while avoiding possible pitfalls. To guide potential researchers, I focus on four issues in this final section: parameters, sources, data collection, and analysis.

Representation and interpretation

My interest in historical representations flows from my epistemological assumptions, which are grounded in post-modernist and post-structuralist thought. 'Reality' is unknowable outside human perception, and there is never only one authority on a given subject. As Friedrich Nietzsche noted, 'There are no facts in themselves. It is always necessary to begin by introducing a meaning in order that there can be a fact' (quoted in Barthes 1981: 15; see also Leander and Neumann in this book). This position does not deny the existence of reality but suggests that the 'true' essence of the object is always unknowable to us. Therefore we must interpret representations of it.

By historical representation, I refer to how the object of an inquiry (X) has been represented over time and space. X can be anything at all: a country (the Congo), a nation or community (the Kurds), a person (Saddam Hussein), or a concept (sovereignty). Societies discursively produce, circulate, and consume representations of X, constructing what are often called 'regimes of truth' or 'knowledge.' These discourses are comprised of signifying sequences that constitute more or less coherent frameworks for what can be said and done. Perhaps the best-known example of this approach is Edward Said's *Orientalism* (1978), in which he exposed how British and French societies constructed 'truth claims' about the supposed innate and inferior qualities of non-white, non-Christian, 'Oriental' people.

Informed by Said and other like-minded scholars, numerous international relations (IR) scholars have studied historical representations. Roxanne Doty's *Imperial Encounters* (1996) compares asymmetrical encounters between Great Britain and colonial Kenya with representations of the Philippines by the United States within its own imperial project. Cynthia Weber's *Simulating Sovereignty* (1995) traces how the meaning of sovereignty has shifted over time within discourses of intervention. Her later book *Faking It* (1999) playfully explores the representation of the Caribbean region in US foreign policy discourses. I will discuss my own work on contested meanings of the Congo in

more detail below. One of my current research projects uses African national parks as its object of inquiry.

All of these studies, and many more, reveal how certain structures of knowledge have been produced and some of their political consequences. Unlike other structural approaches to IR (either Neo-Realist or Marxian), this discursive approach rejects the idea that resources can be explained outside of their discursive context. Rather, social interaction is influenced by cognitive scripts, categories, and rationalities (see Torfing 1999: 81–2). Power is the *practice* of knowledge as a socially constructed system, within which various actors articulate and circulate their representations of ‘truth.’

Since representations of reality and their sequences within discourses are what we work with to understand power, I am primarily concerned with how names, meanings, and characteristics are attached to the world around us. I focus on the *mechanics* of knowledge and identity, and how they differ across time and space. For instance, understanding that this is a ‘tree,’ that is a ‘book,’ and I am a ‘man’ presumes access to commonly shared structures of knowledge about objects such as trees, books, and men. But these naming practices might mean something different (or perhaps nothing at all) to people living in different cultures or historical eras. A tree might be a natural resource to be preserved, a commodity to be harvested, a living soul force to be honored, or an embodiment of the spirits of the dead to be worshipped. So it becomes important to understand that representations are historically and contextually contingent. Specifically, I am looking for the ways that actors represent the object of inquiry. What adjectives, illustrations, or comparisons do they use?

Representations are inventions based on language, but they are not neutral or innocuous signifiers. Because they enable actors to ‘know’ the object and to act upon what they ‘know,’ representations have very real political implications. Certain paths of action become possible within distinct discourses, while other paths become unthinkable. For example, two photos circulated in the media in the aftermath of the August 2005 flooding in New Orleans. The first showed a couple chest-high in the water with bags full of groceries. The caption stated that this couple had ‘found’ food in the wake of Hurricane Katrina. The second photo was of a similar scene, a woman chest-high in the water with a bag full of groceries, but she was identified as a ‘looter.’ This disparity generated much attention because the ‘finders’ were Caucasian while the ‘looter’ was African-American. But even beyond the racial elements at work here, these representations enabled and justified certain actions. Police, for instance, would be expected to assist the couple and to arrest or even

shoot the single woman. Thus discursive practices created a truth-effect – ‘a doing, an activity and a normalized thing in society, one enjoining activity and conformity’ (Brown 2005: 63) – that shaped the possibilities for action. Or, as Iver Neumann states, ‘Because a discourse maintains a degree of regularity in social relations, it produces preconditions for action’ (in this book: 62).

Since some representations become accepted as ‘true’ and others do not, it is important to ask how certain structures of knowledge become dominant. Particular meanings and identities are widely accepted, or ‘fixed,’ not because of any inherent ‘truth’ but because of the strength of that specific representation. The production and circulation of discourses are politically contested, and which discourse will gain social acceptance will depend in large part on the distribution of power (see Leander’s discussion of symbolic power and power relations). Representations are rarely the exclusive product of the object itself, even if it has agency, such as a state or an individual. One must do more than merely examine the utterances of Congolese political leaders or of Saddam Hussein. In both cases, a number of external actors have had greater success in establishing ‘truth’ and ‘knowledge.’

We must ask, Who constructs knowledge and truth claims, for what purposes, and against what resistance? For example, Saddam Hussein may have produced a specific image of himself and his history as Iraqi leader, but his ability to circulate this image and have it gain social currency was limited during his incarceration. In contrast, George W. Bush and his administration had far greater power within the international community to ensure that their representation of Hussein became socially dominant. I suspect that most readers would give little credibility to the representations of Hussein’s identity and history, despite the fact that those discourses have had tremendous salience for Iraqi lives and people elsewhere.

One can investigate the workings of power in the production of discourses by exploring the struggle over who gets to speak authoritatively. External forces are constantly at play, seeking to select, plot, and interpret the events and meanings by which identities are represented. As Said noted, the dominant knowledge of ‘the Orient’ was a creation of the European imperial imagination. With its representations repeated over and over again in Western literature, government publications, and advertisements, Orientalism became *authoritative* knowledge. This helps a researcher disaggregate actors. My research on representations of the Congo led me to investigate the discourses of non-Western actors and, more significantly, forced me to unpack ‘the West’ by focusing on

specific discursive agents and their struggle to gain hegemonic representation. I had to distinguish between Western governments (particularly the United States and Belgium) and explore important divisions within those governments (such as between the White House, CIA, and State Department).

Power is also exercised through the circulation process as competing discourses jockey for greater social acceptance and reproduction. There are often multiple and complex reasons for certain discourses gaining hegemony, and I believe it is important that a researcher be sensitive to these issues. Indeed, while discourses shape power, power also shapes discourse. Power, like discourses, is never totally centralized. A primary goal of this approach is to explore the relationship between discourse and power as they relate to representation (see also Ackerly in this book). The significant points I would underscore here are the multiplicity and contestedness of discourses; the complicated ways in which power works through the production and circulation processes; and the recognition that researchers are not neutral observers, but often are intimately related to the power hierarchies at play.

With regards to agency, this approach assumes that people are guided to act in certain ways, and not others, by their discursively produced understanding of the world and their place in it (see Ringmar 1996). It rejects arguments that actors are motivated by inherent (universal) interests, rational means-ends preferences, or even internalized norms and values. As a fairly macro-level approach, it is admittedly limited in its ability to investigate issues of agency (again, see Leander's employment of Bourdieu). But I am skeptical that micro-level attempts at causal explanation offer better analyses because micro-level analyses usually ignore the effects of discourses as structures of meaning (contrast with Checkel's claims in this book).

So how does one study representations? My own work on the Congo assumed that representational practices are embedded in historical social narratives. Therefore, I combined discourse analysis and historical research to examine struggles over the articulation and circulation of competing narratives. Each of these actors claimed dominant authorship, but obviously, some of these voices were reproduced more than others, giving them greater 'weight.'

Exploring the complexities of this discursive production required me to engage with a wide and diverse spectrum of sources and authors. During the 1960s, for instance, the Congo was rewritten on the floor of the UN General Assembly by representatives from the Soviet Union, newly independent African states (most notably Ghana and Guinea),

Belgium, and the United States, all competing to present their narrative of events. Within the Congo, multiple voices – President, Prime Minister, future coup leader, secessionist leader, local media, citizenry groups, members of the army – articulated either a Congolese national identity or a regional, sub-state identity. Competing narratives also circulated in international and regional media, pamphlets and fliers passed around at political meetings across the globe, government pronouncements from Western and African capitals, best-selling novels, fictional and documentary films, and the ‘bush’ of the Congolese jungle. As I discuss in the next section, I found it necessary to engage in a wide variety of sources when researching, in part to explore the multiplicity and contestedness of discourses, to disaggregate actors, and to explore the complicated ways discourses were circulated and achieved social dominance.

Interpretation requires not just a description of these particular representations and representational practices but a deeper contextualization within the larger structures of meaning of which they are a part. Without going into the theoretical and philosophical debates within the discipline of History (see White 1978; Barthes 1981), let me merely point out that I believe historians produce their own ‘regimes of truth,’ not objective ‘truth.’ History produces its own discourses. Research is highly contested, and the historian is not neutral. This means that both primary and secondary sources should be treated as texts to be decoded and deconstructed. Moreover, this requires a distinction between empiricism as a *method* (skills of verification, close textual attention, proper sourcing, referencing, and so on) and as a *philosophy of knowledge* (the illusion of delivering fact, truth, and a knowable reality). While I (and other ‘post-modern historians’) value the former, we reject the latter.

I find Clifford Geertz’s ‘thick description’ (1973) a useful label for this type of deeply contextualized historical analysis. In particular, I have found the ‘long conversation’ concept of historical anthropologists Jean and John Comaroff to be a useful way of understanding the historic contestation over representations. In their work on the colonial contact between the Tswana peoples of South Africa and the British Christian missionaries, the Comaroffs define the ‘long conversation’ as ‘the actions and interactions that laid the bases of an intelligible colonial discourse’ (1991: 198–9). They argue that there were two faces to this conversation between colonizer and colonized: what was talked *about*; and the struggle to gain mastery over the *terms* of the encounter. I believe that representations are historically produced within similar ‘long conversations,’ where multiple actors come together to contest the meanings of those identities and the terms in which they are expressed.

Drawing from my research on the Congo, one can see such a conversation taking place at the time of Congolese independence. What was under discussion was the extent that the Congolese were 'civilized,' 'developed,' and 'mature' enough to enjoy the 'gift' of independence and sovereignty. One can recognize how various actors struggled to establish both what was talked about and the terms of that conversation.

However, there is a third dimension to the 'long conversation' overlooked by the Comaroffs: the struggle over finding and creating an acceptable position or space within the conversation. Specifically, this refers to the ability to access 'discursive space' within which to engage in the conversation – as Foucault noted, discourses empower certain people to speak (and act). Delineating and policing discursive space has been an important element in international relations, especially for disadvantaged Third World states like the Congo. At times, international discursive space has been actively closed off to competing and counter-hegemonic discourses. For example, immediately after independence, Western governments not only intervened directly to deny the seating of Prime Minister Patrice Lumumba's United Nations delegation, but also his access to the radio station in his country's capital. Both of these actions effectively limited his ability to articulate and circulate his narratives of Congolese identity within the 'long conversation' at the moment of Congolese independence.

Let me reiterate that I am not arguing that the existence or absence of a specific historical representation offers a causal explanation, largely because these representations are historically contingent (see Hermann's discussion of content analysis and Duffy's application of pragmatic analysis, both in this book). For example, the image of Congolese 'inherent savagery' (a familiar Western trope) engendered intervention and colonial conquest in the late 19th century: 'bringing civilization to the savages.' But this same representation enabled Western policies of inaction and indifference to the Congo a century later: 'violence is due to their innate barbarism and tribalism, so there is nothing we can do about it.' Representations do not cause policies, such as intervention, nor do they explain choices, such as whether to intervene at one time rather than another. Representations cannot determine action completely. As Neumann notes, 'Discourse analysis aims at specifying the bandwidth of possible outcomes' (in this book: 62).

I maintain that structures of knowledge establish preconditions and parameters for the possibility of action, rather than explaining why certain choices are made. For example, it helps a researcher understand the range of options imaginable to President John F. Kennedy during

the Cuban Missile Crisis, but it does not explain why he made specific decisions (Weldes 1999). To examine individual decision making, one would need to employ other methods. But while there might be methodological compatibility, one should be sensitive to the possible existence of an epistemological divide on the issue of causality. Personally, I remain unconvinced that we as scholars can offer causal explanations, only reasonably informed conjectures. The world is far too complex and contingent to be studied with any degree of certainty. My post-positivist approach is based on 'a logic of interpretation that acknowledges the improbability of cataloguing, calculating, and specifying "real causes"' (Campbell 1993: 7–8; also see the significant differences between Gusterson and Checkel in this book).

Practical advice for dealing with data

There are several steps to this method, each with its logistical challenges. I will discuss some of these along four general lines: establishing the parameters of a doable project, selecting appropriate sources of data, collecting that data, and analyzing it. But let me preface those comments by pointing out that there is almost always an arbitrary element in case selection (even more than Klotz suggests in this book). Many cases may actually work just as well as the ones you end up choosing. It is always useful to keep in mind that your project should be relevant, enjoyable, and doable.

Simple logistical issues will determine some parameters of your research. For example, basic language limitations matter: if you do not speak or read the language that most of the data is in, you should probably find another case. Or there simply might not be enough information out there to find. But you do have to make others choices for yourself, and you should be honest about why you make them. My comments here aim to help researchers understand the intellectual justifications that underpin the choices involved in historical analysis of representations.

Establishing parameters

It is easy to get overwhelmed by a topic that is just too unwieldy. I find it useful to pick a very narrow, specific topic that allows me to explore much larger issues (note Leander's similar advice). For my dissertation, I chose to examine how the Congo had been represented within the international community, beginning with its colonial conquest up to the current civil war. This case study let me explore not only issues of

colonialism and neo-colonialism, but the social construction of sovereignty, the performativity of stateness, repression, and resistance, and the decline of the Westphalian state system.

However, telling the definitive story of how the Congo has been imagined over the past century would be an overwhelming task, filling numerous volumes. To make my project doable, I focused on four historical moments: the colonial 'invention' of the Congo at the end of the 19th century; its decolonization in 1960; its re-invention as 'Zaire' during the 1970s; and the 'return' of the Congo at the end of the 20th century. (For more guidance on demarcating such historical periods, see Neumann's discussion of 'monuments.')

During each of these four periods, the identity of the Congo was being contested, with numerous forces attempting to produce and attach meanings to its territory and people. These forces sought to create 'regimes of truth' about the Congo by defining and inscribing its identity.

I originally wanted to have six historical moments but found that would require more time and effort than was reasonable. Likewise, I wanted to have one of my historical moments focus on the Ali-Foreman 'Rumble in the Jungle,' and I soon realized that there were a few strong intellectual reasons to include that case beyond it simply being cool – if there had not been a larger justification, being 'cool' would not cut it. When examining historical representations, what matters most is selecting points where forces are seeking to create regimes of truth about the object of inquiry, representation X, by defining and inscribing its meaning. This type of approach stresses historical contingency with a focus on ruptures and disjunctures rather than continuities. In researching the Congo, I chose four moments that seemed to involve the greatest degree of contestation over the Congo's identity and that were historically varied, spanning over a century. Admittedly there can be a bit of arbitrariness to the selection of historical moments, but one should acquire a certain level of background knowledge on the subject in order to identify empirically rich moments of historical rupture.

Sources of data

When I talk about 'data,' I am first and foremost referring to textual representations: attempts to fix the meanings of my object of inquiry, representation X. This tends to be done by numerous actors. Discourse analysis requires employing multiple texts given that 'a single source cannot be claimed to support empirical arguments about discourse as a social background' (Milliken 1999: 233). When researching the construction of Congolese identity, I engaged empirical data from a broad array

of sources. While the majority came from the 'official' realm of governmental reports, speeches, and documents, I also drew from journalism, travel literature, academic treatises, fiction, film, museum displays, art, images, maps, and other 'popular' texts. These texts often provide the most vivid and potent examples of the techniques through which Third World subjects have been narrated by Western hegemonic powers. For many outside observers, including politicians, these are the sources that have provided the primary framework within which the Congo has been made 'knowable.' As David Newbury (1998) pointed out, many Westerners are intellectually uninformed about the Congo but are so inundated by stereotypical images that they feel they have a well-defined cognitive framework. Novels such as *Heart of Darkness*, films such as *Congo*, and cartoons such as *Tintin in the Congo* constitute the basic discursive structure through which many Westerners view the Congo even today.

Different topics will, of course, mean engaging in different sources of data. But I firmly believe in casting the net wide, mainly because our structures of knowledge derive from a variety of sources. Therefore, possible sources include (but are by no means limited to): speeches by political leaders and elites, government records and public announcements, private writings of political elites, popular fiction, non-fiction, newspapers, magazines, music, cartoons, music, television, and the Internet. I will discuss the 'weighting' of various data below, but for now I think it is important to begin with an open mind (see also Ackerly and Neumann). A popular text (that is, a text with wide circulation such as a presidential speech, popular movie, or well-known photograph) will clearly be important in the process of structuring meaning. But more obscure texts (those that have a much more limited circulation, like an academic article or poem by an unknown writer) are often still important, if for no other reason than they represent an alternative to the dominant discourse.

Collecting data

I often combine archival work in historical records with interviews and investigations of popular culture texts. These three sources can each provide their own unique problems. Despite my emphasis on narrowing down potential sources, scarcity of data can also be an issue, since gaining access to data can be challenging.

While I regard the distinction between 'official' and the 'popular' data to be a fiction of the discipline, I employ the distinction here in order to highlight different ways of collecting data for each. The 'official' is

what has traditionally been treated by the political science discipline as 'legitimate' source material: government documents, speeches by state leaders, the writings of political elites, and so forth. What I am calling the 'popular' can be considered the non-traditional: literature, movies, music, cartoons, and so forth. This has generally been designated at the realm of 'popular culture' as opposed to 'political culture.'

The 'official' data relevant for an examination of historical representations are found in a number of places, from libraries and the Internet, but most often in government archives. Without meaning to state the obvious, not all archives are the same. For example, the British National Archive is extremely well organized, with the entire catalog accessible from the Web. But some countries have, shall we say, a different culture about sharing state records. The Belgian archive was very difficult for me to access, and I was denied entry on several occasions. Or it may be that no organized archives exist to house the historical material you are interested in investigating. For example, King Leopold II burned almost all the documents associated with his rule in the Congo immediately before handing control over to the Belgian government. Fortunately, the Belgian foreign ministry had their own copies of many of the torched documents.

Archives in the developing world often are not as organized, accessible, and user-friendly as those in the developed world, possibly for good reasons – ranging from a healthy (and sometimes well-founded) suspicion of Western researchers to neglect and mismanagement to the impoverishment of state infrastructures due to global inequalities. Sometimes state officials might not even be aware of the existence of archives even though they may be in the same building – an experience I have encountered on more than one occasion. It is usually safe to assume that your time in the archive will take longer than you expect. My experience has been that a personal contact at an archive (no matter where it is) is an invaluable asset for the researcher.

Access to popular culture can also be difficult or simply impossible. For example, I have no idea how I would go about accessing texts from Congolese society in the late 18th century. Therefore, I only examined examples of Western fiction and non-fiction writing, from travelogues by colonial explorers and tourists to popular novels by Conrad and Graham Greene. I examined the ways the Congo was discursively represented: As an empty landscape waiting for Western conquest (Stanley)? As a primordial 'heart of darkness' that corrupted civilized Europeans (Conrad)? These were powerful and evocative images that have been re-employed and circulated frequently over time. I also looked at the

representations of the Congo in the popular press. I focused on the major newspapers in Belgium, the United States, and France, including major magazines of the day, such as *Time*, *Life*, and *Newsweek*. I also found it useful to examine how the Congo was portrayed in music, movies, television, and cartoons (a highly fruitful source of data for multiple reasons). Museums, world exhibitions, and public spaces (such as public statues and commemorative arches in Brussels) provided additional rich source material. While by no means a comprehensive sample of how the Congo was portrayed in the Western popular imagination, drawing on the myriad of textual and visual forms by which actors attempt to articulate, circulate, and fix meanings compensates for inevitable limitations in any particular source of data.

A potential limitation is language proficiency. The representations of the Congo exist in numerous languages. For example, there are several major languages in the Congo itself (including French, Lingala, Kiswahili, Kikongo, and Tshiluba), while its colonial ruler, Belgium, has three official languages – none of which are my native tongue. This has meant that countless relevant texts went unstudied by me simply because I could not understand them. And even when I could, I suspect my language skills were not proficient enough to capture subtle meanings, allusions, and jokes. This is a serious problem (see Neumann's observations about 'cultural competency'). Focusing on material only available in your native tongue greatly limits your observations. In the end, I tried to acknowledge these limitations, avoid any overly grand claims, and recognize the narrow focus of my work.

Interviews provide more challenges than I have room here to discuss fully (see Gusterson for elaboration). Gaining access to subjects can often be difficult. Again, language limitations can also be problematic. For instance, I often use an interpreter and rely on him to accurately translate the words and meanings of the speaker, which is often extremely difficult to do. My being a white male also raises gender and racial problems that can often color the exchange, and often in ways that I am unaware. And, of course, interview subjects may simply be untruthful for numerous reasons.

In many cases, the researcher may be faced with data overload, a problem I frequently encountered when doing my Congo research. For example, when investigating historical representations of the Congo at the time of independence, I was simply swamped with what often seemed to be relevant data – from *National Geographic* articles to innumerable political cartoons from the European press to an endless slew of official pronouncements from various governments. If I did not make

hard decisions about what counted and what would not (such as limiting my review of newspapers and news magazines to a handful), I have no doubt I would still be researching today – and in some ways I still am! This gets back to my earlier point about setting parameters: I had to make tough decisions in order to make my project doable, and I had to have solid intellectual reasons for making those decisions. I tried to be as honest and transparent about those decisions as I could (see both Leander and Ackerly on reflexivity). As a result, all my conclusions are tentative and tenuous at best. But I believe that is the nature – and value – of doing qualitative research.

Analyzing data

So what do you look for in the data? Even as I am gathering data, I begin analyzing it. First, I try to identify the different discourses engaged in representing X at a given moment. In what ways do these actors represent the object of inquiry? What type of language do they use when referring to it? For example, at the time of Congolese independence, how did Western leaders in Belgium and the United States portray the country, its inhabitants, and its leaders?

Second, I chart the contestation of these discourses. For instance, why did the Belgian and American presses portray the Congo in different – though equally negative – ways at independence? Who is engaged in the articulation and circulation of these alternative discourses? What is potentially at stake for these actors? Why do certain discourses emerge as socially dominant but others do not? What are the social and political strategies involved in that contestation? How are these discourses being consumed, and by whom?

Third, I historicize and contextualize these representations and discourses within the larger structures of meaning of which they are a part. For example, American representations of the Congo during its independence were situated within a larger Cold War discursive framework, while Belgian representations were part of a longer colonially inspired framework. Sensitivity to history and context allowed me to observe how portrayals of the Congo changed over time. Here is where I also realized how much the ‘official’ sources were informed by ‘popular’ structures of knowledge. During the 1960 crisis surrounding Congolese independence, Western political elites frequently employed texts, metaphors, and images from popular culture, ranging from *Tarzan* movies to Joseph Conrad’s *Heart of Darkness* and H.G. Wells’ *War of the Worlds* to contemporary magazines and cartoons. The reason for this is simple: the structures of knowledge in a society are as much a product of ‘popular’

culture as they are of 'political' culture. The dichotomy between the two is an illusion that obscures more than it reveals.

Finally, I explore how the dominant discourses enable certain policies and practices to become possible. That is, what becomes thinkable and what does not? For example, the Eisenhower Administration's eventual declaration that there was a Congolese problem, and that Lumumba was the source of that problem, had clear political implications: namely, the authorization of his assassination by the CIA. This action was only thinkable because of the representations generated during this time (with their strong historical roots).

Obviously, my approach produced copious amounts of notes (always written on just one side of the page so as to make it easier to find missing quotes or pieces of information later). In this work, I try to track the development of representations and assess their intensity in terms of circulation and social acceptance. I try to structure a narrative of these events – the production, circulation, and contestation of discourses and the range of possible actions they engender. Admittedly, the narrative I produce is an artificial and subjective creation that I use to impose order. Since I am interested in examining historical contingency by focusing on ruptures and disjunctures, I eschew the impulse of traditional historical narratives to portray continuity. In the end, I try to write a convincing narrative that provides an understanding of the 'how' questions which initiated my research.

Conclusion

As I noted at the outset, I do not believe that the world presents itself to us as self-evident. I believe our engagement with it is based on interpretation. As human beings, we make sense of the world around us through the social construction of the meanings, characteristics, and 'truth' that make reality 'knowable.' There is no way to step outside of interpretation. There is no objective Truth to discover, only competing interpretations to navigate.

Since my epistemological position is open to the criticism that it leads to relativism and raises questions about the role of the researcher in the interpretive process, let me respond. I do not believe it is possible to strive for some mythical goal of objectivity, since no such *terra firma* exists. Therefore, I recognize I am not neutral, and I am not too concerned with charges of interpretative bias. But are there ways to decide what counts as 'good' analysis? I believe there are. For me, there

are two important issues to consider when judging the validity of one's interpretation.

First, is there supporting evidence to back up my claims? As a researcher, it will often seem obvious to us that the bulk of our data is pointing to a certain set of interpretations. Of course, our interpretation of that data is what is leading us to our concluding interpretation. But I believe it is important to have supporting evidence. If I claim that the US government portrayed the Congo as Y, which thus enabled it to act in Z manner, I need to provide evidence of both Y and Z. If I cannot, then my claims should be taken as highly speculative. I would argue that this is the reason one needs to do as much historical research as possible. But am I slipping rationality and empiricism back in? I reiterate my distinction between empiricism as *method* versus *philosophy of knowledge*. The value I place on the former does not make my claims 'true,' but it does strengthen my ability to argue for their validity.

This leads to my second point: that the validity of one's interpretation can be measured by its logical coherence does not imply that there is an objective measure of logical coherence (in contrast to a rational choice approach, for instance). Put simply, I am interested in whether or not my conclusions make sense to me, and if they are convincing to others. Do they provide a reasonable answer for the questions I was trying to answer? If not, then I try again. Does such a position lead to relativism? Absolutely. My goal as a researcher is to provide an argument about why my interpretation is valid, so that I can convince others that mine is one of the best interpretations out there. In a very real sense, I am constructing my own representation of the representations I am studying – I am very much part of the process of knowledge construction that I am investigating. Being self-reflexive and honest, I admit that I, like all other researchers, am motivated by an array of personal, political, and intellectual agendas. With my work, I am constructing my own discourses. And because I want them to gain social dominance, I am concerned that my conclusions convince other people.

7

Ethnographic Research

Hugh Gusterson

The anthropologist is always inclined to turn toward the concrete, the particular, the microscopic. We are the miniaturists of the social sciences, painting on Lilliputian canvases with what we take to be delicate strokes. We hope to find in the little what eludes us in the large, to stumble upon general truths while sorting through special cases.

Clifford Geertz (1968: 4)

James Clifford (1997: 56) has, in a much cited locution borrowed from Renato Rosaldo, theorized the methodology of ethnographic research – my craft – as ‘deep hanging out.’ This perverse phrase captures nicely the improvisational quality of fieldwork, the confusing overlap between informal streetcorner conversation and the serious inquiry embodied in ethnographic fieldwork, and the profound level of understanding of the other for which ethnography aims through apparently casual methods.

This phrase ‘deep hanging out’ also hints at a contrast between the methodologies of cultural anthropology (which inclines toward the informal) and political science (which is more tightly buttoned). It is my impression, based on limited observation of the training of graduate students in international relations, that political scientists are expected to go into their dissertation research with well-honed hypotheses that aim to prise open crevices in the existing literature based on a careful parsing of independent and dependent variables and a shrewd selection of case studies that might illuminate the relationships between those variables. Political Science graduate students often seem to know what their dissertation will argue, and what the chapter outline will look like, before they have got deeply into the research. While anthropology

graduate students spend years acquiring language skills, working on pre-dissertation literature reviews and writing dissertation proposals, these proposals often focus more on broad questions suggested by the existing literature than on hypotheses to be tested. Meanwhile dissertation committees in anthropology departments tend to expect student research plans to shift as they encounter the vicissitudes of the fieldwork environment: bureaucratic difficulties in accessing a particular site, research subjects disinclined to discuss the topic that seemed so crucial in the student's literature review, research subjects passionately interested in discussing issues the student had not thought to inquire about, and unpredicted events (riots, protests, scandals, conflicts, funerals, celebrations, and so on) that provide unforeseen but compelling windows onto an unfamiliar cultural world.

Moreover, although there are stories of anthropologists such as Melville Herskovits insisting that his students mail their fieldnotes to him from the field for review, most anthropologists report that they received minimal guidance about fieldwork from advisers and dissertation committees either before they went to the field or while they were there. I myself, for example, have never seen another anthropologist's notes, and I am far from unique in that regard (see Sanjek 1990). Anthropologists often assume that each fieldwork situation is different, and that researchers will have to improvise accordingly. Furthermore, first fieldwork is a 'rite of passage' (turning graduate students into mature anthropologists), and it is part of the ritual testing to throw students on their own resources.

In this chapter, stressing the simultaneous informality and rigor of ethnographic fieldwork, I shall take the reader through the key components of 'the ethnographic method.' Although anthropologists often use methods that overlap with those of other disciplines – archival research, written questionnaires, and formal interviews, for example – I focus here on methodological concerns more unique to the ethnographic encounter: gaining access to the field; doing semi-structured interviews and what anthropologists oxymoronically refer to as 'participant observation'; navigating the ethical obligations of fieldwork; and writing up research first through fieldnotes and, later, in ethnographies. (The word 'ethnography,' confusingly, refers both to a method of research and to the finished literary product.) Until the upheavals in anthropology of the late 1980s and 1990s, anthropologists were most likely to study non-Western cultures rather than Western, metropolitan cultures; to study a single localized site; and to focus their studies on those subordinate in status. Recent years,

by contrast, have seen the increasing legitimacy in the anthropology of 'repatriated anthropology,' 'multi-sited ethnography,' and 'studying up' (Nader 1974; Clifford and Marcus 1986; Marcus and Fischer 1986; Marcus 1995).

I shall draw opportunistically on the relatively small methods literature in anthropology and on what I know of others' fieldwork, but I shall also draw considerably on my own experience doing ethnographic research among American nuclear weapons scientists and, to a lesser extent, antinuclear activists. My original dissertation fieldwork in the San Francisco Bay Area in the late 1980s, part of the disciplinary transformation, was on the Lawrence Livermore National Laboratory, the nuclear weapons laboratory that designed the warheads for missiles (Gusterson 1996). I was trying to understand how scientists came to feel that they had a vocation to design nuclear weapons; I also wanted to describe the phenomenology of weapons work, the effect of weapons work on marital and family relationships, the relationship between the weapons laboratory and local institutions ranging from churches to the town council, and the impact upon the laboratory of the sizeable antinuclear protests of the early 1980s. I interviewed many of the protestors as well, and at one point accompanied a group of anarchists from the Bay Area on a weeklong protest trip to the Nevada Nuclear Test Site.

More recently, for a follow-up book, I have been doing multi-sited fieldwork among weapons scientists at both the Livermore and the Los Alamos nuclear weapons laboratories; among antinuclear activists in California, New Mexico, and Washington DC; and sporadic interviews with senior bureaucrats from the nuclear weapons complex wherever I can find them. If my earlier fieldwork focused largely on rank-and-file weapons scientists, this research has been more centered on senior managers of the weapons laboratories and on major players in the Washington defense bureaucracy – busy decision-makers who are not easily accessed. The purpose of this research is to trace the process by which the national security bureaucracy (especially the nuclear weapons complex) came to acquiesce in the suspension of nuclear testing and the negotiation of the Comprehensive Test Ban Treaty in the early 1990s (Gusterson 2004). If the first research project was anchored to a single, localized site – the Livermore Laboratory – the second project has, in keeping with a more general anthropological evolution away from a preoccupation with the local, focused much more on diffuse networks, structures, and processes that are both national and international in scale.

Accessing the field

Like the space shuttle entering the earth's atmosphere, the ethnographer entering the field must get the angle of approach just right, or the resultant friction may burn up the mission. Unlike shuttle astronauts, ethnographers have widely varying missions, each with different optimal angles of approach. Sometimes what opens the village doors can be quite unpredictable, especially to an outsider. Paul Stoller (1989: 40–1) reports that he made little headway in penetrating the world of sorkos – magician-healers in Niger – until the day a bird defecated on his head. This was taken by a sorko who witnessed it as sign that Stoller was chosen for apprenticeship.

What works for one ethnographer seeking entrée to the field may prove disastrous for another. Margaret Harrell (2003), for example, is an anthropologist who studied US military families. She reports that a letter from a commanding officer directing military personnel to cooperate with her was indispensable to her fieldwork. By contrast, the anthropologist Philippe Bourgois (1995), who did fieldwork with crack dealers in New York's Spanish Harlem, would have been crippled by the endorsement of uniformed authorities and, in his case, being mistreated by the police on one occasion helped his fieldwork considerably. In general, ethnographers entering the field seek to ally with gatekeepers who will vouch for them and to avoid falling in with the wrong crowd – the only problem being that, as you enter an unfamiliar cultural situation, it can be quite hard to tell which is which.

Ethnographers are inevitably marked in the field by their race, class, gender, education level, nationality, and other characteristics. In some contexts, aspects of the researcher's own identity may play a facilitating role; in others they may be crippling. It is hard, for example, to imagine a woman doing Loic Wacquant's (2003) research with boxers in Chicago, or a man doing Elizabeth Fernea's (1969) research among the wives of a sheikh in Iraq or Stephanie Kane's work with female prostitutes (1998). Ethnographers inevitably have to decide which aspects of a field environment are more or less accessible or closed off by virtue of their own identity.

In my own case, when I decided to do an ethnography of the Lawrence Livermore National Laboratory, my problem was that I was a foreign (British) citizen attempting to study a top secret military facility where I knew nobody and to which access was largely forbidden for those without clearances. I thought of making a formal approach to the Laboratory's management for permission to study the facility, but

decided the likelihood was low that such permission would be granted and, once denied, it was not inconceivable that Lab management would actively obstruct more informal approaches to their weapons scientists. In the end, I tried a scattershot approach of three simultaneous entry strategies, only one of which was truly fruitful and one of which was nearly quite damaging.

The first strategy, joining my practical need for accommodation with my interest in meeting laboratory employees, was to look for a room in a house occupied by lab employees. Over the course of 2 years of field-work, I lived in three different houses with different kinds of laboratory employees – a technician, a computer programmer, and an engineer. Over time, I heard a lot of gossip about the Lab from these employees, who I got to know well as individual friends. However, they did not introduce me to many other lab employees, and it is dangerous to rely on single sources to understand a complex institution employing over 8000 people. I felt as if I were slowly developing a deep understanding of very tiny and isolated pockets of laboratory life from my roommates.

Roommates were, however, a particularly good source of basic orientation information. Disorientation is one of the strongest sensations of the ethnographer newly arrived in the field. Consequently, the beginning of field research is often dominated by an attempt to simply get one's bearings by asking lots of very basic questions. In my case, these questions included the following: Why do some people have red and others green badges? How many directorates are there at the Lab, and what do they all do? What is that tall building in the middle of the Lab I can see from the perimeter? What kinds of clothes do people wear to work at the Lab and how should I dress when meeting them? Is it alright to talk about 'bombs' or should I call them 'devices'? What is a CAIN booth? (It regulates access to restricted areas of the Lab for those with clearances. An employee stands in the booth and swipes a card, as if at an ATM, entering a secret code, and is then granted admission.)

My second strategy was to make use of one chance contact I had made at a party a few weeks before coming to Livermore. At this party I met a woman and her husband, who worked as a scientist at the Lab. They both lived in Livermore, and the wife was especially interested in my research. She invited me to lunch with a promise that she would provide me entrée to a wider network in Livermore. I noted that she brought her teenage daughter to lunch and seemed uncomfortable. I was fortunate to discover from a friend of the couple that her husband (who was not keen on talking to me) was concerned that my interest in his wife was not purely academic, so I moved on. As my research unfolded, I

observed that scientists were happy to talk to me when I was introduced through networks of scientists at the Lab, but often resisted talking to me if the introduction came through their spouses.

The approach that worked, my third strategy, was the result of extraordinary serendipity. My graduate student advisor mentioned to me that he was supervising an undergraduate thesis on the town of Livermore by a student who grew up there. I contacted the student and found that his father worked at the Lab. The son arranged for me to go and visit his father. I anticipated discussing with the father the feasibility of my study and getting his advice on how to approach people. Instead, when I arrived at his home at seven o'clock one evening, he said, 'Take out your notebook. I will tell you my life story.' I said very little for the next 2 hours, at the end of which I had pages of fascinating material about a man who had fled North Korea as a teenager, come to the United States with nothing, trained as a physicist, and sought work as a weapons scientist because of what he referred to throughout the interview as his 'monolithic anticommunism.' He demonstrated for me that evening that the way to understand lab employees was not to ask a series of abstract questions about their ideological beliefs but to elicit life histories that crystallized their commitments in narratives of the events through which they were enacted – a technique whose power has been beautifully demonstrated in Faye Ginsburg's (1989) ethnography of pro-life and pro-choice activists in the Mid-West, published just 2 years after my conversation with the Korean scientist in Livermore. At the end of our encounter, the scientist offered to put me in touch with five more lab scientists if they agreed. They did. Each of them referred me to still more colleagues, and the rest was history.

This technique of building an exponentially increasing network of research subjects from an original subject zero is referred to in the methodological literature, for obvious reasons, as the 'snowball technique.' Its strength is that people who trust one another trust those referred to them through the network. Its weakness is that it does not operate through random sampling, and there is an obvious danger that the ethnographer will get trapped inside the network's echo chamber and will be confused by what he or she hears there for the wider discourse of an entire institutional setting (see also Ackerly in this book). In my own case, I was confident that I was reaching a wider sample partly because my collection of interviewees was so large, and partly because I deliberately pushed interviewees to refer me to others chosen to diversify my sample.

I also, over time, further diversified my pool of subjects by searching for interlocutors in other settings too. In a context where about three quarters of weapons scientists identified themselves as active Christians, church attendance proved an important way of getting to meet them as well as building relationships with their pastors, who also became interview subjects. I joined a softball and a basketball team at the Lab; I joined the Lab singles group (more of a Friday evening and weekend outings club than a dating arrangement); I hung around bars in town, and I sometimes went for lunch to the Lab cafeteria, which was open to the public and proved a good place to cajole scientists I already knew into introducing me to others.

The pool of interlocutors I developed through these techniques has been important also for my newer research on the weapons laboratories' adaptation to the end of nuclear testing. I have gone back to some of these interlocutors to explore their reaction to life in a weapons laboratory without nuclear testing. However, my new research has focused much more on very busy senior managers than my earlier research did. In securing interview access, I have been fortunate to be able to build on the success of the first research project: that research secured me a professorship at MIT, which is a highly respected institution at Livermore and Los Alamos. Senior managers there will usually make time to talk to an MIT professor. Beyond that, my original research has now been widely profiled in local newspapers, it produced a book that many lab employees have read, and I have written a number of articles for local newspapers. This has given me a measure of legitimacy around town, and it gives potential interlocutors a sense that I am a known quantity who can be trusted as much as any outsider can. One lesson to draw is that when anthropologists' relationships to research sites carry on across a decade or more, as they often do, they deepen over time, opening up new vistas of understanding.

Participant observation

If you asked an older generation of anthropologists to define 'the ethnographic method,' they would put 'participant observation' at the center of it. Participant observation, the essence of the 'deep hanging out,' denotes a method of research in which ethnographers join in the flow of daily life while also taking notes on it (either in real time or shortly afterwards). If the locals went hunting, harvesting, drinking, feasting, or pilgrimaging, the anthropologist tried to go with

them, often to do it with them, and to record as accurately as possible what was said and done.

There are many obvious benefits to participant observation. First, this level of sustained contact with research subjects helps to build relationships of trust and intimacy with them. Second, seeing for oneself what people do and choosing what to record of it is surely far better than learning about it after the fact in a fragmentary fashion from documents or informant interviews. It is the difference between sitting in someone's living room with them and peeking in through a keyhole. Finally, participant observation is a particularly effective way of exploring the difference between the 'frontstage' and 'backstage' – between formal, idealized accounts of a culture and the messy divergences of actual practice. Imagine what a Martian ethnographer would believe about the way an American university works if they relied on formal interviews with faculty and staff, and then imagine what they would learn instead if they went to faculty meetings and gossipy lunches with the staff while living in a student dorm in the evenings, and you will get my point.

Some of my favorite ethnographies use participant observation for particularly good effect. In *Peyote Hunt*, Barbara Myerhoff (1976) accompanies a group of Huichol Indians led by Ramon, a shamanic figure, on a long pilgrimage into the Mexico desert to the Huichols' original mythic home and home still to their gods. Their pilgrimage culminates with the sacred ingestion of peyote and with the harvesting of the hallucinogenic buttons for rituals for the coming year. Her participation in the pilgrimage and its visionary culmination enables her to get inside Huichol cosmology and mystical religious experience as much as any outsider can. Myerhoff's narrative has a cinematic quality. As she relates, with a novelist's eye for detail and drama, the pilgrims' jokes, the reader feels that he or she is alongside the Huichols in their journey. *Of Two Minds*, by Tanya Luhrmann (2001), looks at the socialization of American psychotherapists and psychiatrists. Her description of the way medical residents learn their trade and internalize diagnostic categories of mental illness is particularly enlivened by the fact that she put herself through the same apprenticeship in order better to understand it.

Participant observation has been especially important in ethnographic investigations of American poverty. This is because there is often a sharp divergence between, on the one hand, judgmental assumptions about the poor that circulate in the media and among policy makers and, on the other hand, the lived experience of poverty. In books such as Carol Stack's *All Our Kin* (1997) and Philippe Bourgois' *In Search Of Respect* (1995), privileged white ethnographers reposition themselves by living

in the midst of poor black and Hispanic communities. More effectively than any dry, statistics-laden policy study, these ethnographies build a picture of the exhausting daily grind of lives lived in poverty, of creative adaptations to poverty that are also entrapping (such as crack dealing), and of the barriers to escaping the ghetto that are so much more clearly visible from within than outside. But the ultimate exercise in participant observation in poverty was conducted not by a professional ethnographer but by the journalistic public intellectual Barbara Ehrenreich. In her justly celebrated book, *Nickel and Dime* (2002), she goes undercover, working as a low-end waitress, a hotel maid, and a Walmart worker. Ehrenreich records not only the mass of petty brutalities against the poor in the workplace but also keeps an exact ledger of the financial costs faced by low-income workers versus the income they can secure. By the end of the book, one thinks it a miracle anyone moves up from this life at all.

Given the insights participant observation facilitates, I regret the limited role it played in my own fieldwork among weapons scientists. Although I spent as much time as possible simply 'hanging out' with Lab employees in church, in their homes, and on hikes, I sometimes wonder what I might have seen had I been allowed to come into the Lab day after day with my notebook and fade into the background. Anthropologists of science who have been given full access to scientific laboratories have often written ethnographies that focus on the micro-processes through which scientific facts are constructed (Latour and Woolgar 1986; Fujimura 1996; Knorr Cetina 1999). I suspect that, had I engaged in participant observation within the Lab itself, I would have written an ethnography more focused on disputes over weapons design details, the bureaucratic relationships between different ranks and categories of employees, and the phenomenological disconnect between small daily tasks within the laboratory and the laboratory's larger project of developing a massive arsenal of weapons of mass destruction capable of liquidating hundreds of millions of people. As it was, my enforced positioning on the margins of laboratory life produced an ethnography that foregrounded secrecy practices within the Laboratory and the Laboratory's relationships with other institutions, and my residence outside the Laboratory fence but within the homes of weapons scientists made me particularly sensitive to the role and experience of laboratory spouses.

Despite the circumscribed role participant observation played in my field research, there are still things I would not know without having engaged in it. For example, I recall being in the cafeteria of the Livermore

Laboratory when CNN started to broadcast the story of the Oklahoma City bombing. As I watched weapons scientists around me turn up the volume on the cafeteria TV and, using CNN's details about blast damage, rush to calculate the power of Timothy McVeigh's bomb on the backs of their white paper table napkins, I viscerally understood something about the phenomenology of their craft. Other informal interactions have also been instructive. By befriending a new Lab employee and watching her mounting anxiety as her investigation for a security clearance dragged on for months, I came to understand, better than I could through interviews, the indispensability of a clearance, the petty humiliations of life without a clearance, and the terror an employee feels at the prospect of denial. Taking a long and beautiful dog-walk with another employee, I was stunned by a torrent of criticism of the Director of Los Alamos that he (and his colleagues) had held back in other interactions. I have also found that rank-and-file weapons scientists talking over a beer joke to the detriment of their managers and evince much more skepticism about the new simulation technologies being developed at the weapons labs than similar scientists do in tape recorded interviews or than managers in any context I can access. Rank-and-file weapons designers' informal narratives of the origin of these simulation technologies are more likely to stress pork barrel deals in Washington, whereas more formal interviews with managers accent the scientific and technical logic of the technologies and the overall rationality of the program of stockpile stewardship. In other words, my ability to 'hang out' with ordinary weapons scientists gave me special insight into the gulf between 'frontstage' and 'backstage' narratives of the stockpile stewardship program, between what is said in public and what is whispered or said jokingly in private.

A second example comes from my parallel fieldwork among antinuclear activists. As these activists prepared to go on a week-long protest to the Nevada Nuclear Test Site, I attended their preparatory workshops where I heard first-hand about people using sick days and vacation time to keep their jobs while they went on the protest. I joined with them as they role-played being arrested and subjected to police brutality, and having their planning meetings infiltrated by undercover police officers. Then I traveled with them to the Nevada Desert, where I lived in a tent for a week with no running water and was taught by those around me how to deal with the extremes of heat and cold in the desert in spring. Finally, I shared their experience of civil disobedience. Without having gone through all this myself, I do not think I could so easily grasp the extraordinary sense of community among the activists, the sacrifices many of the protestors made to be there, or

the relationship between the privations of protest and the strange rush of euphoria from civil disobedience. Nor, without my time amongst the protestors, would I have recognized with such clarity the mistaken nature of comments made by members of the Livermore community characterizing the protestors as communists and unemployed folks who had nothing better to do.

A final contribution made by participant observation is more amorphous and mysterious, but no less important for that. It concerns the reformation of my own emotional relationship to nuclear weapons. When I arrived in Livermore in the mid-1980s, I did so as someone who had been deeply concerned about the possibility of superpower nuclear war to the point of even having occasional nightmares about it. By the time I left Livermore 2 years later, I had lost my subjective fear of nuclear weapons and have never been able to recover it. It just disappeared! I am unable to give a precise account of the processes involved here but it is clear that, in some way, living amongst people who joked about nuclear weapons and took for granted the human ability to control these weapons, I absorbed their sense of ease – or, if you prefer, their ability to live in denial.

Semi-structured interviews

The core of my research consisted of semi-structured interviews organized around the elicitation of life histories. I collected well over a hundred of these interviews, which were almost always tape-recorded. This was important because I was interested in the exact language scientists used to describe their beliefs and experiences, and because my interlocutors attached great importance to precise quotation of their remarks. In my original research it was through such semi-structured interviews that I came to understand how weapons scientists understood the ethics and politics of their work, how they reconciled their weapons work with their religious commitments, how they experienced the weapons design process emotionally, and how weapons work affected family life. In my more recent research, I have used such interviews to reconstruct negotiations about the end of nuclear testing at the higher levels of the weapons bureaucracy, to understand the purpose of new simulation technologies being built at the weapons labs, and to elicit the response of rank-and-file weapons scientists to the end of nuclear testing and the emergence of virtual nuclear weapons science.

Many social scientists, less interpretively focused than I, are deeply concerned about the exact comparability of their subjects under the

research microscope. Sociologists devising questionnaires, for example, seek to ensure that, however diverse their pool of research subjects, they are responding to the same questions. Here it is the consistency of the questions posed to different individuals or populations that enables the sociologist to make differentiating generalizations: everything comes back to the way different people respond to the same questions. If each interview or questionnaire is different, then comparison is clouded.

While the benefits of such a research protocol are obvious, it also acts as a straitjacket. If, as Sharon Hutchinson (1996) says, ethnography is 'the fine art of conversation,' individuals like to talk about different things and, by insisting on precise comparability, this research methodology prevents the detailed exploration of individuality. It also tends to bore research subjects, forcing them into a kind of mass-produced superficiality. In my interviews there was a core set of questions I asked everyone: where were you educated? To what level and in what subject? What are your religious commitments? What is your work at the Lab or in the antinuclear movement? How did you come to decide to do weapons work? Has anyone in your family or beyond given you a hard time for working on weapons? Such questions, as well as producing a matrix for comparison, served as icebreakers and orienting probes for deeper conversations that followed. But beyond this elementary set of common questions, my interviews with different research subjects diverged quite substantially as I followed strategies I call 'branching' and 'building.'

My interviews followed a 'branching' pattern as I tailored them to individual interests and identities. Interviews followed different trajectories for physicists and engineers, for the elite weapons designers and the scientists who worked under them, for Christians, Jews, and atheists. Interviews also branched in different directions as my line of questioning responded to what individual scientists showed particular interest in discussing.

As for 'building,' each interview built upon earlier ones as my understanding of the Lab deepened and expanded over time, and interviews I did at the end of the research project were quite different from those conducted at the outset. I came to think of myself as having conversations not just with unique individuals, each fascinating in his or her own right, but also with a single entity: a discourse community. As these unfolding conversations suggested recurrent discursive themes, new avenues of inquiry, or newly evident lacunae in my own understanding, so the questioning shifted, each conversation establishing

a new beachhead as I probed more deeply into the culture of the Lab or, sometimes, circled back recursively to check anomalies and uncertainties.

Researchers who subscribe to more positivist understandings of the world than I do assume that research subjects have stable 'values,' 'preferences,' 'beliefs,' 'ideologies,' or 'cultures' and that it is the researcher's job to find out what they are as cleanly as possible (to some extent, Checkel and Hermann, in this book). But I soon noticed that subjects I interviewed more than once might contradict themselves in interesting ways, or that some interviewees presented themselves quite differently to journalists and to me. Positivists would see such fluctuations as 'noise' to be eliminated in order to ascertain what the informant 'really' thinks. I came, instead, to see these instabilities of discourse as themselves part of informants' cultural identities. And if, for example, a scientist's statements about the Russians showed little fluctuation while his or her comments about the ethics of weapons work were variable, this variability was itself an important ethnographic datum.

Just as Lao Tzu said that no two stones can be thrown in the same river, so I would say that it is not possible to interview the same subject twice. Thus, rather than thinking that I was sampling or eliciting a stable, pre-existing reality as objectively as possible, I began to think of interviews as dynamic events through which the identity of the subject was performed and even co-constructed by the interviewer and interviewee. In these conversations, interviewees did not so much manifest an unchanging essence there, like some geological pattern, plain for any researcher to see if they knew how to scrape away the surface. Instead they drew on the complex repertoires of their speech community to perform themselves in response to particular lines of questioning (How is your work ethical? Do you think nuclear war will happen? How do you deal with antinuclear activists?) that often reflected my own past in the antinuclear movement. A different interlocutor with different preoccupations would have provoked different performances of self since, as Renato Rosaldo (1989: 19) observes, 'the ethnographer, as a positioned subject, grasps certain human phenomena better than others. He or she occupies a position or structural location and observes with a particular angle of vision.' And, of course, as my earlier discussion of the way my interviews built upon one another makes clear, I was changed by each interview too: no two interviews were done by the same interviewer.

At their worst, these interviews produced the ethnographic equivalent of American Presidential debates: stale performances using rehearsed lines and recycled snippets from the Laboratory's public relations

campaigns. At their best, the interviews produced performances of self in modest kitchens and living rooms around Livermore that were profound, touching, revelatory, funny, counter-intuitive, and educative. The role of interviewer affords a license to ask questions of a kind that would not normally be permitted for strangers – indeed, even for friends in most contexts – while the act of sustained, attentive, supportive listening can be powerfully enabling for the person being heard (and indeed, in a different way, for the listener as well). This kind of listening – accompanied by requests to clarify apparent contradictions, to tie emotions to recalled events, or to address narrative gaps – can induce a creatively reflective state of mind as interviewer and interviewee move into a zone of interaction that hybridizes therapeutic encounters and journalistic interrogations.

Some of my interviews lasted 4 hours. One lasted for 15 hours, spread over a series of sessions, which a retired scientist taped as a bequest for his daughter. (When I attended his funeral after he died of Alzheimers a few years later, I felt a secret and special bond to him.) I began to realize that, as scientists reflected on the ethics of their work, reconstructed their decisions to come to the laboratory, and recalled their emotional responses to nuclear tests they had experienced, they were sometimes opening spaces they shared with few others. One wife, eavesdropping on my interview with her husband, interrupted to say, ‘How come you told him that? You’ve never told me that!’ Many scientists told me that they thought about the ethics of their work but none of their colleagues did – a clear indication that everyone was thinking about nuclear ethics, but quietly and in private. The interviews, then, generated articulations not only of fiercely public ideologies, but also of the private, the whispered, the half crystallized on the edge of consciousness. And once these articulations became public, as they were pushed back into the community through my writing, then in a modest way they changed the field of discourse I had come to study.

Inscriptions

In earlier generations, anthropologists passed many of their evening hours typing up index cards. These cards enabled them to store and sort information they had gathered on, say, patrilineal cross cousins, funeral rituals, or witchcraft beliefs. Doing fieldwork in the computer age, I use the cut-and-paste function of Word to do some of the work for which those anthropologists used index cards. However, I mainly organize my notes around interviews and interactions with individuals, recording

their exact words whenever possible. The fundamental organizing principle of my notes, then, is the individual biography, though I do also sort information on my hard drive and in manila folders around themes. Sometimes I take a pair of scissors to printed transcripts of interviews, scattering textual shards to differently themed manila folders. Clearly there is a relationship between the organization of my notes around individuals and the fact that my writing often makes use of long quotes from individual informants and, on occasion, features extensive profiles of individual research subjects (Gusterson 1995a,b).

How do ethnographers know when it is time to leave the field and start writing? Often, they have no choice: their research funds dry up or their sabbaticals end, and they go home with whatever notes they have. In my own case, I felt that fieldwork was getting stale when I found myself often able to predict how research subjects would answer my questions. While I was still learning new things, this meant that my understanding of the culture was achieving a certain depth and stability and was, to some degree, plateauing. It was time to stop talking and start writing.

In preparing to write, I read my notes on interviews with individual interlocutors, as well as transcriptions of them, flagging recurrent patterns, variations on themes, and quotable passages. The recurrent patterns have ranged from noting that Livermore scientists are more optimistic about simulation technologies than Los Alamos scientists to observing the use of similar metaphors by different people, often people who do not know one another. Examples include the use of birth metaphors to describe the process of designing and testing a nuclear weapon, the use of machine metaphors to describe the human body, and the use of anthropomorphic metaphors to describe machines.

Ethnographers of my generation, often influenced directly or indirectly by the writings of Michel Foucault, tend to see human cultural worlds as constructed by the intersecting power of ingrained cultural practices and the discourses through which people speak about their world. When we do fieldwork we note these practices and we record as much of the discourse as we can, looking for recurrent patterns. Just as psychotherapists have to talk to people at the conscious level in order to deduce what is happening in their unconscious worlds, so anthropologists have to observe and talk to individuals (or groups of individuals), but are really interested in the practices and discourses that transcend the level of the individual and, to put it in Foucauldian terms, provide the social material from which their individuality is constructed.

(See Neumann and Dunn, in this book, for examples of finding such discursive patterns at other levels of analysis.)

Writing this up as ethnography poses its own set of challenges since, compared to political science monographs, the criteria for writing and judging ethnography are much looser and more heterogeneous. Ward Goodenough (1981), likening culture to language, said that an ethnography was a sort of cultural grammar book, and that, just as a grammar book would teach you how to speak a language, so an ethnography should teach you how to behave appropriately in a particular culture. Clifford Geertz (1973), seeing culture as a text to be interpreted rather than a set of rules to be followed, thought good ethnography gave deep insight into the nuances of an alien lifeworld and into the meanings by which its adherents struggled to live. Carolyn Nordstrom (2004: 14) meant something similar when she said 'ethnography must be able to bring a people and a place to life in the eyes and hearts of those who have not been there.'

Such descriptions imply that the goal of ethnography is just particularistic description. However, as the quote by Clifford Geertz at the front of this chapter makes clear, the point of ethnography is to describe the particular in order to illuminate 'general truths' – the functioning of capitalism, the nature of ritual, the experience of oppression, say. In general, anthropologists would agree that good ethnography gives a rich evocation of the cultural world it describes while also contributing something to theory and being of interest to those who are not specialists on the culture area described. They would also say that it should 'feel right' to those other area specialists – though not being without surprises – and that it should give a thick enough description that readers could draw their own inferences about the culture being described.

The index card generation of anthropologists often said in their writing that 'the data suggest' and 'it was observed that...', but who collected these data and by whom was it observed? References to 'data' and the use of the passive voice – the hallmarks of what Donna Haraway (1988) calls the 'God's eye view from nowhere' – are the familiar tropes of phony objectivism. Of course, we check our facts, quote people accurately, and do our best to make sure we know what people mean by what they say, but in the end 'data' are collected and written up by individual researchers who have their own concerns, insights, and blind spots. For this reason, as I have done throughout this chapter, I use the first person in my writing, in order to remind the reader that the 'data' have been collected, sifted, organized, and represented by a particular

individual who readers have to decide whether or not they trust. As a way of helping readers to make up their minds, at the end of my first book, *Nuclear Rites*, I also gave a page each to a handful of key informants to comment on the book.

Human subjects and ethics

In the United States, government agencies such as the National Science Foundation (NSF) and the National Institutes of Health (NIH) require that research they fund be approved by university panels for the protection of human subjects and refuse to disburse money until these review boards have approved it. In the wake of scandals such as the death of Jessie Gelsinger, a healthy 18-year-old killed in 1999 by poorly conceived gene therapy research at the University of Pennsylvania, universities are also increasingly concerned to review the safety of human subjects in research conducted by their students or faculty (Stolberg 1999). (For an example of a human subjects tutorial and exam, see <http://web.mit.edu/committees/couhes/>.) While the process of human subjects review gives universities more control over research for which they may be legally liable, it can also benefit researchers, since the university effectively legitimates the research it has approved and indemnifies researchers in the event of legal action.

Many anthropologists see human subjects review boards as, at best, institutions that slow research with unnecessary red tape and, at worst, the preserve of curmudgeonly bureaucrats from other disciplines who do not understand the unique exigencies of ethnographic fieldwork. In the past, conflicts have focused in particular on consent forms. Human subjects bureaucracies like consent forms because they clarify the contract between researchers and subjects while providing tangible evidence that subjects agreed to be studied. Anthropologists often dislike consent forms, first, because their subjects may not be able to read and are often suspicious of people bearing bureaucratic paperwork and, second, because in many Third World countries (especially those with overly energetic police forces) the quickest way to lose a subject's friendship and cooperation is to ask them to sign a form saying they agree to inform on their country to a foreigner. Consequently, anthropologists are sometimes tempted to engage in research under the human subjects bureaucracy radar or to diverge from written protocols in research practice.

Readers should not infer from this that anthropologists are indifferent to the well-being of their subjects. In my experience, the opposite is true.

But, in keeping with the informality of anthropology, it is often assumed that human subjects are best protected not by inflexible bureaucratic codes but by ethnographers who think situationally about an internalized mandate to 'do no harm.' Such a perspective is affirmed by the current language in the American Anthropological Association (AAA) ethics code (<http://www.aaanet.org/committees/ethics/ethcode.htm>) (See also Fluehr-Lobban 1998, 2003)), which states,

[I]t is understood that the informed consent process is dynamic and continuous; the process should be initiated in the project design and continue through implementation by way of dialogue and negotiation with those studied. Researchers are responsible for identifying and complying with the various informed consent codes, laws and regulations affecting their projects. Informed consent, for the purposes of this code, does not necessarily imply or require a particular written or signed form. It is the quality of the consent, not the format, that is relevant.

The 1971 version of the AAA ethics code took a particularly strong stance against secret consulting by ethnographers. Reflecting general disapproval of anthropologists who secretly consulted for the American national security state during the Vietnam War, it said,

'In accordance with the Association's general position on clandestine and secret research, no reports should be provided to sponsors that are not also available to the general public and, where practicable, to the population studied ... Anthropologists should not communicate findings secretly to some and withhold them from others.'

In response to lobbying from anthropologists who consult for the private sector and are concerned about proprietary data, that language has now been watered down. The current AAA ethics code merely says that anthropologists 'must be open about the purpose(s), potential impacts, and source(s) of support for research projects with funders, colleagues, persons studied or providing information, and with relevant parties affected by the research.'

Still, even in its contemporary weakened version, the ethics code stresses the importance of obtaining the informed consent of those being studied:

'Anthropological researchers should obtain in advance the informed consent of persons being studied, providing information, owning or

controlling access to material being studied, or otherwise identified as having interests which might be impacted by the research.'

This is quite different from the ethics code of, say, the American Psychological Association, which allows for the routine deception of subjects in psychological experiments, provided this deception has been approved by human subjects review boards and as long as it is explained to research subjects after the completion of the experiment.

Two famous scandals in anthropology underline the ethical dangers of the ethnographic method. In 1983, Stanford University (the department in which I was trained) denied a PhD to Steven Moshier on ethical grounds. Among the concerns, he was accused of taking photographs without their consent of women undergoing abortions and of endangering research subjects who criticized China's birth control policies by not concealing their identities (Sun 1983; Turner 1983; Lee 1986). And the journalist Patrick Tierney (2000) unleashed the biggest controversy in 30 years by claiming that, in the 1960s, James Neel had exacerbated a deadly measles epidemic among the Yanomami of Venezuela through his inappropriate use of a flawed vaccine and that Napoleon Chagnon, complicit with Neel, staged fights among the Yanomami to make his documentary films more interesting, among other charges. In the confusing debate that followed, Tierney softened some of his allegations, and over time the charges against Neel began to look much weaker than those against Chagnon (Sahlins 2000; Borofsky 2005).

Such scandals aside, most anthropologists do show concern for the well-being of the human subjects with whom they work. If one listens to corridor talk among anthropologists, they tend to be concerned about protecting the confidentiality of their interlocutors and about advocating for underprivileged communities they study. Many anthropologists donate book royalties or other income to communities with whom they may have a lifelong research relationship, and they often go to special lengths to secure medical or educational help for individual interlocutors with whom they have particularly close relationships. One of my colleagues at MIT recently paid for the medical care of an ailing informant, for example, and then for his funeral.

Anthropologists who work in war-torn parts of the world also fret that their work might inadvertently facilitate government repression, the maneuvers of death squads, and so on. It is said, for example, that some anthropological work on Mayan textile patterns may have helped Guatemalan death squads identify indigenous communities for liquidation. The French anthropologist Georges Condominas was horrified to learn that the US government had (illegally) translated and distributed

his ethnography of a Vietnamese people to Green Berets during the Vietnam War and that his research subjects were subsequently tortured (see Berreman 1980). There are even instances of anthropologists who have left book manuscripts unpublished out of such concerns. Ever since the AAA was torn apart in 1968 by revelations that some anthropologists were secretly consulting on counter-insurgency in Southeast Asia for the US national security state, most anthropologists have kept their distance from such agencies as the CIA, the Department of Defense, and even USAID that might be interested in their knowledge of populations around the world (Berreman 1974; Wakin 1992; Price 2000, 2004). After 9/11, some suggested that anthropologists should contribute their expertise to the war on terror by working more closely with US national security agencies, but this suggestion has been more condemned than approved within anthropology (Gusterson 2003, 2005; Wax 2003; McFate 2005; Moos 2005a,b; Price and Gusterson 2005).

As for my own research, I have had to make sure that my interlocutors understood why I was interested in talking to them. Most of them had PhDs and worked in bureaucratic contexts; they were reassured by a consent form stating that a university Institutional Review Board (IRB) had approved my research, that I was funded by a well-known foundation, and that set forth the contractual terms of our conversations. The most reassuring of these contractual terms was that I promised not to quote them by name – an easy commitment for me to make since it is conventional for anthropologists to invent pseudonyms for those they portray in their writing. The only exception I have made to this rule has been for very senior officials in the weapons bureaucracy who are often quoted in the newspaper and who give explicit permission to be quoted by name.

There were three respects in which my fieldwork relationship with human subjects was unusual for an anthropologist. First, most of the people I interviewed had top-secret clearances and I had to take special care not to jeopardize those clearances. In some cases that has meant not using information people have shared in indiscreet moments; in others it has meant taking particular care to obscure the source of information that, whether or not it is officially secret, does not usually circulate in the public sphere. Second, the antinuclear activists I studied are subjects not only of my inquiring gaze but also, often, of government surveillance. I have been acutely aware that it is difficult to draw a clear line between writing that explains the cultural logic of the antinuclear movement in ways activists themselves might appreciate and writing that might feed into the intelligence-gathering of government agencies that do not wish these activists well. I have tried to write about the symbolic and ideological systems of activists rather than about their operational

procedures, though this skews my writing on this subject. Third, my commitment to fieldwork among both weapons scientists and activists – two communities deeply antagonistic to one another – poses a special burden. I have had to make sure that each community understood that I was also talking to their antagonists, but also to take care not to let either community use me as an intelligence agent against the other.

Conclusion

At the outset, I emphasized that ethnographic methods are simultaneously rigorous, informal, and improvisational. There is, obviously, a tension between these three descriptors, but I believe it is a creative one. While I have benefited enormously from reading the work on my research specialty, nuclear politics, and culture, by scholars from other disciplines, I am struck that no other research methodology enables the investigator to grapple with the lived experience of people in the way that ethnography does. Historians are confined by the documents they can find or by the decades-old memories of interviewees; psychologists only access the minds of their subjects through questionnaires or highly staged interviews; while political scientists often reify their material through the deployment, unpersuasive and metaphysical to this analyst, of assumptions about the rational calculations of human actors or the methodological separability of so-called ‘dependent’ and ‘independent’ variables. Ethnography is always in danger of lapsing into memoir or journalism at one extreme or obscuring the human beings it studies with relentless theorization at the other, but its creative stew of investigative techniques also holds the promise of a human(e) science that seeks objectivity without objectifying its subjects, that balances rigor with reflexivity, and understands that human action cannot be investigated apart from the local meanings attached to it.

Acknowledgment

My thanks for helpful feedback to Audie Klotz and the students in her graduate seminar on methods at Syracuse University. I also received valuable comments from my MIT colleagues Rufus Helmreich, Stefan Helmreich, Jean Jackson, Philip Loring, Heather Paxson, Susan Silbey, Susan Slyomovics, and Chris Walley. Special thanks to Heather and Stefan for excavating information about the Steven Mosher case and for inquiring into the archeology (or is it genealogy?) of the phrase ‘deep hanging out.’

8

Process Tracing

Jeffrey T. Checkel

'This argument is too structural. It's under-determined and based on unrealistic assumptions. Moreover, it tells us little about how the world really works.' Among many scholars – the present author included – this is an oft-heard set of complaints. Consider two examples. The central thesis of the democratic peace literature – that democracies do not fight other democracies – is hailed as one of the few law-like propositions in international relations. Yet, as critics rightly stress, we know amazingly little about the mechanisms generating such peaceful relations (Rosato 2003: 585–6, *passim*; Forum 2005; Hamberg 2005). And scholars have for years debated the identity-shaping effects of European institutions. One claim is that bureaucrats 'go native' in Brussels, adopting European values at the expense of national ones. Yet, here too, critics correctly note that we know virtually nothing about the process and mechanisms underlying these potentially transformative dynamics (Checkel 2005a,b).

So, to paraphrase a former American president, 'it's the process stupid.' To invoke process is synonymous with an understanding of theories as based on causal mechanisms. To study such mechanisms, we must employ a method of process tracing. Process tracers, I argue, are well placed to move us beyond unproductive 'either/or' meta-theoretical debates to empirical applications where *both* agents and structures matter. Moreover, to capture such dynamic interactions, these scholars must be epistemologically plural – employing *both* positivist and post-positivist methodological lenses.

But realizing this epistemological–methodological promise is not easy. Proponents of process tracing should be wary of losing sight of the big picture, be aware of the method's significant data requirements, and recognize epistemological assumptions inherent in its application. To

develop these arguments, I provide first the basics of a process- and mechanism-based approach to the study of international politics. The next section draws upon my own experience as an inveterate process tracer to outline how the technique works in practice. I then assess the method and, finally, conclude with several reflections on the epistemological challenges of this focus – challenges that should push process tracers to evince a new level of pluralism in their work.

Causal mechanisms and process tracing

Mechanisms operate at an analytical level below that of a more encompassing theory; they increase the theory's credibility by rendering more fine-grained explanations (Johnson 2002: 230–1). According to one widely cited definition, a mechanism is 'a set of hypotheses that could be the explanation for some social phenomenon, the explanation being in terms of interactions between individuals and other individuals, or between individuals and some social aggregate' (Hedstroem and Swedberg 1998: 25, 32–3; see also Hovi 2004). As 'recurrent processes linking specified initial conditions and a specific outcome' (Mayntz 2003: 4–5), mechanisms connect things.

For example, in a recent project on international socialization (Checkel 2005a,b), our objective was to minimize the lag between international institutions (cause) and socializing outcomes (effect) at the state or unit level. To this end, I theorized three generic social mechanisms – strategic calculation, role playing, and normative suasion – which allowed me to posit more fine-grained connections between institutions and changes in state interests and identities.

How does one then study these causal mechanisms in action? Process tracing would seem to be the answer as it identifies a causal chain that links independent and dependent variables (George and Bennett 2005: 206–7; Odell 2006: 37–8). Methodologically, process tracing provides the how-we-come-to-know nuts and bolts for mechanism-based accounts of social change. But it also directs one to trace the process in a very specific, theoretically informed way. The researcher looks for a series of theoretically predicted intermediate steps.

Conceptually, when talking of mechanisms and process tracing in this chapter, I have adopted a micro-perspective. Theoretically, this means I examine what are sometimes called 'agent-to-agent' mechanisms (George and Bennett 2005: 145). Empirically, I focus on specific decision-making dynamics (see also Hermann and Post in this book). However, this is merely a pragmatic choice, not an ontological claim.

I know this micro-level best, theoretically and empirically. Not all mechanisms need to be linked to individual decisions. Others have argued for a macro-focus in the study of causal mechanisms (Tilly 2001; Katzenstein and Sil 2005). Whether the specific lessons I offer can be scaled up to a more macro-level is a question for future research. Epistemologically, process tracing is compatible with a positivist or, to be more precise, scientific realist understanding of causation in linear terms.

In sum, process tracing means to trace the operation of the causal mechanism(s) at work in a given situation. One carefully maps the process, exploring the extent to which it coincides with prior, theoretically derived expectations about the workings of the mechanism. The data for process tracing is overwhelmingly qualitative in nature, and may include historical memoirs, expert surveys, interviews, press accounts, and documents (see Gheciu 2005a,b for an excellent application). Process tracing is strong on questions of interactions; it is much weaker at establishing structural context. Logistically, the greatest challenge is the significant amount of time and data that it requires.

In principle, process tracing is compatible with, and complementary to a range of other methods within the empiricist/positivist tradition. These include statistical techniques, analytic narratives (Bates *et al.* 1998), formal modeling (Hoffmann in this book), case studies (Klotz in this book), and content analysis (Hermann in this book). Process tracing is utilized by both empirically oriented rational-choice scholars (Schimmelfennig 2005) and conventional constructivists (Lewis 2005).

Process tracing in action: the case of European institutions

To illustrate this micro-level process tracing tool kit, I assess the causal impact of international socialization. In Europe, there are numerous tantalizing hints of such dynamics, for example, in the EU's Convention on the Future of Europe (Magnetite 2004) or in the European Commission (Hooghe 2005). There are also ongoing, contentious, and unresolved policy disputes (*Economist* 2002, 2003) and academic debates (Laffan 1998; Wessels 1998) over the extent to which European institutions socialize – that is, promote preference and identity shifts. Moreover, with its thickly institutionalized regional environment and a supranational, polity-in-the-making like the EU, Europe seems a most likely case for socialization to occur (Weber 1994; Zürn and Checkel 2005).

Socialization refers to the process of inducting new actors into the norms, rules, and ways of behavior of a given community. Its end point

is internalization, where the community norms and rules become taken for granted (Checkel 2005a). One way to reach this end point is via persuasion, which I define as a social process of communication that involves changing beliefs, attitudes, or behavior, in the absence of overt coercion. It entails convincing someone through argument and principled debate (Zimbardo and Leippe 1991; Perloff 1993: 14; Brody *et al.* 1996; Keohane 2001: 2, 10). To employ my earlier language, it is a social mechanism where the interactions between individuals may lead to changes in interests or even identities.

Persuasion may thus sometimes change people's minds, acting as a motor and mechanism of socialization. However, the key word is 'sometimes.' The challenge has been to articulate the scope conditions under which this is likely to happen. Deductively drawing upon insights from social psychology (Orbell *et al.* 1988) as well as Habermasian social theory, recent work suggests that persuasion (and its close conceptual relative, arguing) is more likely to change the interests of social agents and lead to internalization when: (H1) the target of the socialization attempt is in a novel and uncertain environment and thus cognitively motivated to analyze new information; (H2) the target has few prior, ingrained beliefs that are inconsistent with the socializing agency's message; (H3) the socializing agency/individual is an authoritative member of the in-group to which the target belongs or wants to belong; (H4) the socializing agency/individual does not lecture or demand, but, instead, acts out principles of serious deliberative argument; and (H5) the agency/target interaction occurs in less politicized and more insulated, in-camera settings (see Checkel 2005a for details).

This theorizing – done before I began my research – structured everything that followed. Given that persuasion was the causal mechanism whose effects I sought to explain, process tracing was the obvious methodological choice for studying it. How I studied persuasion and the kinds of data I needed to collect were dictated by these five hypotheses. Specifically, H1 and H2 required detailed knowledge of the target, his/her background, and beliefs on the subject at hand. In a similar fashion, for H3, I needed to collect data on the individual/agency doing the socializing – and especially his/her perceived status. Interviews were crucial for gathering these kinds of data; I then used secondary sources (media appearances, memoirs) as a supplement.

For H4 and H5, the data collection was more demanding as these hypotheses capture the interaction context of the attempt at socialization. In my case, this context was a series of committee meetings in an international organization. Obvious data sources would be

interviews with committee members and minutes of the committee meetings. If the latter are unavailable, interviews with the secretary or administrative person in charge of the committee's operations would be a second-best proxy.

With my theory, hypotheses and ideal data sources now specified, I turn to the example: process tracing socialization dynamics in the Council of Europe as it debated issues of citizenship and nationality in the early and mid-1990s. The Council is a pan-European organization whose mandate is human rights. When it confronts a new issue, it sets up committees of experts, composed of representatives from Council member states as well as academic and policy specialists. Their mandate is to think big in an open way. In the early 1990s, two such committees were established: a Committee of Experts on National Minorities and a Committee of Experts on Nationality. If new norms were these committees' outputs, then the issue for me was the process leading to such outcomes. In particular, what role was played by persuasion?

For the committee on national minorities, there were few attempts at persuasion throughout its five-year life. Rather, committee members were content to horse-trade on the basis of fixed positions and preferences. Key in explaining this outcome was the politicization of its work at a very early stage (H5). Events in the broader public arena (the Bosnian tragedy) and within the committee led to a quick hardening of positions. These political facts greatly diminished the likelihood that the committee's formal brainstorming mandate might lead to successful acts of persuasion, where Council member states might rethink basic preferences on minority policies.

The story was quite different in the committee on nationality. Through the mid-1990s, nationality was a rather hum-drum, boring issue, especially compared with the highly emotive one of minorities. Initially, much of the committee's proceedings were taken up with mundane discussions of how and whether to streamline immigration procedures and regulations. In this technical and largely depoliticized atmosphere, brainstorming and attempts at persuasion were evident, especially in a working group of the committee. In this smaller setting, individuals freely exchanged views on the meaning of nationality in a post-national Europe. They sought to persuade and change attitudes, using the force of example, logical argumentation, and the personal self-esteem in which one persuader was held. In at least two cases, individuals did rethink their views on nationality in a fundamental way, that is, they were convinced to view the issue in a new light (Checkel 2003).

That last sentence, however, raises an important methodological issue. How does this tracing of the process allow me plausibly to assert a

causal role for persuasion as a mechanism of socialization? Put more prosaically, how would I recognize persuasion if it were to walk through the door? I employed multiple data streams, consisting of interviews with committee members (five rounds spread over 5 years), confidential meeting summaries of nearly all the committee's meetings and various secondary sources, and triangulated across them (see also Pouliot 2007: 19; Dunn in this book).

In the interviews, I asked two types of questions. A first touched upon an individual's own thought processes and possibly changing preferences. A second was more intersubjective, asking the interviewee to classify his/her interaction context – a step dictated by hypotheses H4 and H5 above. I gave them four possibilities – coercion, bargaining, persuasion/arguing, and imitation – and asked for a rank ordering. Interviewees were also asked if their ranking changed over time and, if so, why (Checkel 2003).

These methodological injunctions aside, how did I really know that two individuals 'did rethink their views on nationality in a fundamental way?' How did I know these two were persuaded, and not strategically dissimulating or simply emulating others? I began with before and after interviews of the two individuals concerned – that is, interviews just as the committee started to meet and then again after one of its last sessions. I asked specific questions of their views on nationality, why they held them, if those views had changed, why they had changed, and what role(s) coercion, bargaining, persuasion, or imitation had played in the process.

Of course, one should never simply rely on what people say, so I triangulated. This meant that I cross-checked the story related by the two interviewees with other sources. The latter included interviews with other individuals who had observed the first two in action and an analysis of the committee's meeting minutes. The latter are typically not verbatim transcripts; moreover, they are highly political documents as committee members must approve their content before release. Members could – and did – have items (attributions of particular views, say) deleted from the summaries. To mitigate this (potentially huge) source of bias, I took the additional step of interviewing and getting to know the committee secretary, whose responsibility was to write up the minutes.

Collecting data in this theoretically informed way allowed me to reconstruct committee deliberations, building a plausible case that: (a) the views of the two individuals concerned had indeed changed; and (b) that persuasion (as opposed to imitation or bargaining) was the motor driving such shifts. I then further bolstered this claim – derived

from my process tracing – by asking the counterfactual: absent these persuasive dynamics, would the outcome have been any different? In fact, the regional norms to emerge from the committee's deliberations were different from what otherwise would have been the case. For example, on the question of dual nationality, a long-standing prohibitory norm was relaxed, thus making European policies more open to the possibility of individuals holding two citizenships (Council of Europe 1997, 2000).

Finally, moving outside the bounds of the case summarized above, my findings are consistent with insights drawn from laboratory experiments in social psychology on the so-called contact hypothesis (Beyers 2005) and from work on epistemic communities in IR theory (Haas 1992). Of course, ultimately, one can 'never know' as we are not privy to private thought processes. However, the step-wise, cross-checking procedure outlined here sharply bounds and minimizes the danger of erroneous inference.

Assessing process tracing: the good, the bad and the ugly

What have I learned from more than a decade of using process tracing as my method of choice? I offer 12 lessons – four *good*, five *bad*, and three *ugly*. The *good* is the value added that comes from applying the method – how it advances the state of the art methodologically, theoretically, and meta-theoretically. The *bad* are issues and failings of which to be aware before starting a research project with this method. The *ugly* stand out as 'red flags' – questions in need of attention. Addressing the latter will require process tracers to transgress both meta-theoretical (*agents and structures*) and epistemological (*positivism and post-positivism*) boundaries. In discussing the lessons within each category (*good, bad, and ugly*), I proceed from the practical (method) to the conceptual (theory) to the philosophical (meta-theory).

Lesson #1 (Good – Method): coming to grips with first mover advantages

Process tracing can minimize the problems of the so-called first mover advantage (Caporaso *et al.* 2003b: 27–8). If they are honest, most scholars will admit to having favorite theories. In empirical research, the tendency is first to interpret and explain the data through the lens of this favored argument. By encouraging researchers to consider alternative explanations, the positivist-empiricist tool kit has built-in checks against this first mover advantage. And process tracing can make such

checks stronger. Predicting intermediate steps between independent and dependent variables essentially produces a series of mini-checks, constantly pushing the researcher to think hard about the connection (or lack thereof) between expected patterns and what the data say.

Lesson #2 (Good – Method): answering ‘how much data is enough?’

Process tracing makes it easier to address a question that often plagues qualitative researchers: ‘When is there enough data?’ My work on socialization in European institutions provides a case in point. After two rounds of interviewing, I took a break from data collection. Writing up the results – connecting the data to the causal story I was attempting to tell – allowed me to see where my data coverage was still weak. This suggested the kinds of data I would need to collect during future field work. Especially with interviews, I employed what is sometimes called a branching and building strategy, where the results of early interviews are used to restructure and refocus the types of questions asked at later points (see also Gusterson in this book).

After two more rounds of field work, I again wrote up the results, seeking ‘to fill in the blanks’ in my causal-process story. This time, I also circulated the draft to several colleagues. Based on their input and my own, I came to a determination that I had indeed collected enough data. More specifically, I felt that my story was now plausible in that a rigorous but fair-minded reviewer would read the analysis and say ‘yeah, I see the argument; Checkel has made a case for it’ (see also Dunn on establishing valid interpretations).

Lesson #3 (Good – Theory): helping to bring mechanisms back in

A very diverse set of social theorists now call for more attention to mechanisms (compare Elster 1998; Wendt 1999: ch. 2; Johnson 2006). There are good and sensible reasons for this trend. Most important, it moves us away from correlational arguments and as-if styles of reasoning toward theories that capture and explain the world as it really works. Less appreciated are the methodological implications. Simply put, if one is going to invoke the philosophy-of-science language of mechanisms, then process tracing is the logically necessary method for exploring them (see also Drezner 2006: 35).

Lesson #4 (Good – Theory): promoting bridge building

Process tracing has a central role to play in contemporary debates over theoretical bridge building (see Adler 1997). To make connections

between different theoretical tool kits – rational choice and social constructivism, most prominently – scholars have advanced arguments on temporal sequencing and domains of application. Implicitly or explicitly, the method on offer is typically process tracing, as it is extremely useful for teasing out the more fine-grained distinctions and connections between alternative theoretical schools (Fearon and Wendt 2002; Caporaso *et al.* 2003a; Kelley 2004; Checkel 2005b).

Lesson #5 (Bad – Method): proxies are a pain

Process tracers often decry the unrealistic proxies that quantitative researchers employ in the construction of data sets (for example, Hug and Koenig 2000, 2002). But qualitative researchers, including process tracers, face similar problems, albeit at a different level. A central concern in my own work has been to theorize and document the causal mechanisms of socialization, such as persuasion. Did I ever actually see somebody persuaded? Did I see a decision-maker change his or her mind? No, I did not. I was not a fly on the wall, secretly observing these individuals. Participant observation was not an option. I, too, was therefore forced to rely on proxies – before and after interviews, documentary records of the meetings, and the like. At an early point, the process tracing, qualitative scholar thus needs to think hard about the conceptual variables at play in his/her project, and ask what are feasible and justifiable proxies for measuring them.

Lesson #6 (Bad – Method): it takes (lots of) time

Process tracing is time intensive and, to put it ever so delicately, ‘can require enormous amounts of information’ (George and Bennett 2005: 223). Researchers need to think carefully about their own financial limits and temporal constraints. My studies of socialization included five rounds of interviews spread over 5 years and a close reading of numerous documents (both public and confidential). In large part because of its methodology, the project has taken a long time to bring to fruition. While all scholars face trade offs when thinking about productivity, research endeavors, and methods, these dilemmas may be particularly acute for process tracers.

Lesson #7 (Bad – Theory): just how micro to go?

Process tracing and the study of causal mechanisms raise a difficult ‘stopping point’ issue. When does inquiry into such mechanisms stop? How micro should we go? In my project on socialization, I took one mechanism – socialization – and broke it into three sub-mechanisms:

strategic calculation, role playing, and persuasion (Checkel 2005a). Why stop at this point? Persuasion, for example, could be further broken down into its own sub-mechanisms, most likely various types of cognitive processes. My justification has two parts, neither of which has anything to do with process tracing. First, the state of disciplinary knowledge told me that it was a concept like socialization – and not persuasion – that was ripe for disaggregation into smaller component mechanisms (see also Alderson 2001). Second, a growing and increasingly sophisticated array of qualitative techniques (cognitive mapping, interview protocols, surveys) made it possible for me to craft reliable proxies to measure persuasion's causal effect (see also Johnston 2001, 2007).

Lesson #8 (Bad – Theory): non-parsimonious theories

Process tracing is not conducive to the development of parsimonious or generalizable theories (see also Drezner 2006: 35). In part, the reasons for this are social theoretic. As I argued earlier, process tracing is synonymous with a mechanism-based approach to theory development, which, as Elster correctly argues, is 'intermediate between laws and descriptions' (Elster 1998: 45). However, in equal part, the reasons are human and idiosyncratic. The typical process tracer is a scholar driven by empirical puzzles. He/she is happy to combine a bit of this and a bit of that, the goal being to explain more completely the outcome at hand. The end result is partial, middle-range theory (George and Bennett 2005: 7–8, 216). If one is not careful, middle-range theory can lead to over-determined and, in the worse case, 'kitchen-sink' arguments where everything matters. Early attention to research design can minimize such problems (Johnston 2005).

Lesson #9 (Bad – Theory): missing causal complexity

Like any method, process tracing abstracts from and simplifies the real world – probably less than many others, but abstract it still does. By tracing a number of intermediate steps, the method pushes a researcher to think hard about the role played or not played by a particular mechanism. Yet in many cases, the outcome observed is the result of multiple mechanisms interacting over time. Process tracing can help deal with this challenge of causal complexity, as can creative applications of agent-based modeling (Hoffmann in this book; see also Cederman 2003: 146). For instance, process tracing helped me establish when persuasion was present and when it was absent. The latter 'non-finding' then suggested a role for additional causal mechanisms, such as bargaining (Zürn and Checkel 2005: 1052–4).

Lesson #10 (Ugly – Meta-Theory): losing the big picture

In making a methodological choice to examine questions of process, it is all too easy to lose sight of broader structural context. For example, when I presented my findings on individual decision-makers and the social-psychological and institutional factors that might lead them to change their minds in light of persuasive appeals, interpretative scholars noted that I had no way – theoretically or methodologically – for figuring out what counted as a serious deliberative argument. I had just assumed it adhered to the individual, but it was equally plausible that my persuader's arguments were legitimated by the broader social discourse in which he/she was embedded. In positivist–empiricist terms, I had a potential problem of omitted variable bias, while, for interpretivists, the issue was one of missing the broader forces that enable and make possible human agency (compare Neumann and Dunn in this book).

There are two ways of responding to such a problem. One is to deny its validity, along the lines of 'Nobody can do everything; I had to start somewhere.' A second is to view such problems – and their resolution – as a chance to promote genuine epistemological and methodological pluralism within the community of process tracers, a point to which I return below.

Lesson #11 (Ugly – Meta-Theory): losing the ethics

Process tracers may be particularly prone to overlook normative-ethical context. In my collaborative project on socialization and European regional institutions, all participants adopted a mechanism-based approach, and many combined this with a process tracing method (Gheciu 2005a; Lewis 2005; Schimmelfennig 2005). Yet, while we were tracing such dynamics, we forgot to ask important normative-ethical questions. Is it legitimate and just that West Europe – through the EU, NATO, and the Council of Europe – imposes norms and rules on applicant countries from East Europe that in some cases (minority rights) are flagrantly violated by those very same West European states? What are the implications for democratic and legitimate governance if state agents acquire supranational allegiances and loyalties?

Lesson #12 (Ugly – Meta-Theory): the dreaded 'E' word

Most process tracers are empirically oriented scholars who just want to conduct research on the fascinating world around us. On the whole, this is a healthy attitude. Especially for rational-choice scholars who adopt process tracing (Schimmelfennig 2003; Kelley 2004), variable-oriented

language fits well with their positivist–empiricist epistemological orientation. But constructivist theorists are split, with some explicitly (Wendt 1999: 82; also George and Bennett 2005: 206) or implicitly (Ruggie 1998: 94) endorsing the method, while others appear much more skeptical (Adler 2002: 109). Still others advocate a so-called bracketing strategy for capturing such dynamics (Finnemore 1996).

Yet, it is unclear if process tracing in general or bracketing as a specific strategy for implementing it are consistent with the mutual constitution and recursivity at the heart of constructivist social theory (see also Pouliot 2007). Process tracing only works if you hold things constant in a series of steps: A causes B; B then causes C; C then causes D; and so on. Bracketing means, first, to hold structure constant and explore agency's causal role, and, then, to reverse the order, holding agency constant while examining structure's role. These are very linear processes. Indeed, those interpretative constructivists who do employ process tracing are careful to separate it from the discursive and narrative techniques at the heart of their approach (Hopf 2002).

To (begin to) address this state of affairs, the dreaded 'E' word must be revisited. As some have noted (Zehfuss 2002: chs 1, 6; Guzzini 2000), constructivists – and especially those who endorse methods like process tracing – do need more carefully to explicate their epistemological assumptions. And such a rethink will likely require a turn to post-positivist philosophies of science.

Conclusion

After the numerous criticisms in the preceding section, readers may be surprised by my bottom line: Process tracing is a fundamentally important method – one that places theory and data in close proximity (see also Hall 2003). One quickly comes to see what works and – equally important – what does not. This said, process tracers need to think harder about the logical and philosophical bases of this mechanism-based approach. Positivism as a philosophy of science will not do the trick, given its correlational view of causation, instrumental use of theoretical concepts, and narrow methodological writ (Wight 2002).

One possible post-positivist starting point would be scientific realism, which is the 'view that the objects of scientific theories are objects that exist independently of investigators' minds and that the theoretical terms of their theories indeed refer to real objects in the world' (Chernoff 2005: 41; see also Wendt 1999: ch. 2; George and Bennett 2005: 147–8, 214). For many scientific realists, these 'real objects'

are precisely the causal mechanisms of the process tracing studies highlighted in this chapter.

Scientific realism is also inherently plural in that ‘no one method, or epistemology could be expected to fit all cases’ (Wight 2002: 36; more generally, see Lane 1996). With such qualities, it would seem ideally placed to give process tracing conceptual grounding, and allow process tracers not just to triangulate at the level of methods, but across epistemologies as well. Indeed, my own decade-long, hands-on experience as a process tracer suggests that if we want to offer better answers to the questions we ask (Lesson #10 above), then such epistemological and methodological boundary crossing is both essential *and* possible (see also Hopf 2002; and the excellent discussion in Pouliot 2007).

Given such foundations, process tracers can then begin to ask hard questions about their community standards – standards anchored in a philosophically coherent and plural base. What counts as a good mechanism-based explanation of social change and what counts as good process tracing? How can discourse/textual and process tracing approaches be combined?

Building upon but going beyond – epistemologically – the ‘process tracing best practices’ advocated by Bennett and Elman (2007: 183), I would argue that good process tracing adhere to the following core maxims.

- *Philosophy*: It should be grounded, explicitly and self-consciously, in a philosophical base that is methodologically plural, such as that provided by scientific realism or other post-positivist epistemologies, including analytic eclecticism (Katzenstein and Sil 2005), pragmatism (Cochran 2002; Johnson 2006), or conventionalism (Chernoff 2002, 2005), for example.
- *Context*: It will utilize this pluralism both to reconstruct carefully causal processes and to not lose sight of broader structural–discursive–ethical context.
- *Methodology I*: It will develop and carefully justify a set of proxies that will be used to infer the presence of one or more causal mechanisms.
- *Methodology II*: It will take equifinality seriously, which means to consider the alternative paths through which the outcome of interest might have occurred.

While positivists have avoided such issues by focusing excessively on correlation and design at the expense of causation and method (King *et al.* 1994; see also Drezner 2006: 35; Johnson 2006), too

many interpretivists for too long have simply sidestepped methodological questions altogether (Checkel 2006; Hopf 2007). The goal ought to be to give IR process tracers a middle-ground philosophy and epistemology that can fill the vast methodological space between positivism and post-structuralism. This chapter, the edited book of which it is a part, and other recent endeavors (Lebow and Lichbach 2007) hold out the promise of correcting this truly odd state of affairs.

Acknowledgment

For comments on earlier drafts, I am grateful to Hannah Allerdice, Petra Hejnova, Dick Price, the students of Political Science 694 at Syracuse University (Fall 2005), and, especially, Audie Klotz and Deepa Prakash.

This page intentionally left blank

Part III: Boundary Crossing Techniques

This page intentionally left blank

9

Political Personality Profiling

Jerrold M. Post

The Political Personality Profile was developed in order to provide senior policy makers with a comprehensive psychological representation of leaders in context. It describes the life course that shaped key attitudes, and specifies aspects of behavior especially relevant to policy makers dealing with leaders in summit meetings and other high-level negotiations, as well as in crisis situations. The approach asks two general questions: What were the events and experiences that helped shape the leader's personality (psychogenesis)? And what are the psychological forces within a personality that drive political behavior (psychodynamics)? We never can know for certain what drives an individual, but the more solidly we understand these foundations of the leader's psychology, the more confidently we can infer influences on – and patterns of – political behavior.

The Political Personality Profile characterizes the leader's core political personality. With its emphasis on the life course, it integrates longitudinal and cross-sectional analyses. In addition to traditional elements of clinical psychological assessment, elements of the profile include management style, negotiating style, strategic decision-making, crisis decision-making, rhetorical style, cognitive style, and leadership style. By combining them and specifying the political context in which the leader is operating, the Political Personality Profile produces a fuller picture, which identifies how the leader's core personality influences these important leadership characteristics. Another major difference is that the clinician interviews the subject directly, whereas leader assessments are developed indirectly. The interview of individuals who have met personally with the subject has been found to be extremely valuable in remedying this shortfall, and by interviewing a number of informants, one can reduce the likelihood of observer bias. In my profile of

Saddam Hussein, for instance, I was able to interview, either directly or by telephone, six former diplomats and business executives who had had personal contact with him.

The term 'personality' connotes a systematic pattern of functioning that is consistent and coherent over a range of behaviors and over time. To identify these deeply ingrained patterns, it is essential to integrate the life experiences that gave form to that political personality. But not all political situations engage the political personality, so the Political Personality Profile also seeks to identify which political issues are especially salient. The task is to identify which issues 'hook' the leader's political personality, but in so doing one must always analyze the leader in context. We must go beyond the family environment to encompass the historical, political, and cultural context as well.

Since the method I will be describing modifies the perspective of clinical psychology and psychiatry, the aspiring profiler may well ask: does this mean I have to be a psychologist or a psychiatrist to employ this method? The answer is distinctly No. I have been teaching the craft for some 35 years now, and most of the students have a background in political science and history, having taken few or no psychology courses. What is necessary is psychological mindedness. What I teach is the manner in which the life course influences the development of personality and how to draw psychological inferences from behavioral observations.

Let me give an example. A number of years ago, I was asked to develop a profile of the president of a Latin American country. In developing the psychobiographic study of the subject, his prior academic career was a rich source of psychologically relevant material. The Latin American president under my lens, during his academic career, had written extensively, and I had read all of his writings that I could get my hands on. In his major work, even his footnotes had footnotes, and there was an introductory note to one chapter, which was quite remarkable: 'The reader is advised to skip this chapter. It is dull and boring to the extreme. But for the sake of completeness, I must include it.' This punctilious concern for thorough scholarship to the point of perfectionism, even to the point of tedium, had alerted me to pursue behavioral observations that might confirm my hunch that the president had significant compulsive features in his personality. On my visit to the capital city to debrief key informants, I met one afternoon with the deputy chief of the US mission (the DCM), who had met frequently with the president under study. When I interviewed him, he was still fuming. The notoriously gridlocked traffic was particularly bad that day, and although he

had left in what he thought was plenty of time for his one o'clock appointment with the president, he arrived at 1:03, by which time the agitated president had already called the embassy to complain about the DCM's lateness.

Primed by my working hypothesis that the subject of the study was probably quite compulsive, and struck by the exaggerated emphasis on punctuality in a country where time was usually treated very casually, my first question was, 'Describe his desk.' 'Interesting you would ask that,' the DCM responded. 'The president had two neat piles in front of him, which he kept straightening to ensure they were perpendicular to the desk's edge. And, in the midst of our conversation, he looked at the wall, leapt up, saying, "Excuse me, but that picture is tilted," and proceeded to straighten out a picture on the wall that was no more than a quarter of an inch out of kilter.'

The combination of punctilious scholarship, identified in the prior psychobiographic research, and the president's emphasis on punctuality and neatness suggested to me significant compulsive features in his personality, which could have important implications for negotiations. He could be expected to be conscientious and live up to his commitments. Moreover, his written words probably could be taken to reflect his dominant political goals, including the probability of nationalizing his nation's natural resources.

Most personality assessment systems attempt to assess three different dimensions: cognition, affect or feelings, and interpersonal relations. Different theorists will emphasize different dimensions, but understanding all three dimensions is important to addressing political personality. One distinction between the Political Personality Profiling method and the other profiling methods is the emphasis on psychogenesis and psychodynamics; it seeks to answer questions about both events that shape personality and psychological forces within personality that drive political behavior. The framework described is neither strictly Freudian, Jungian, Adlerian, nor Lacanian; it strives to help the policy consumer understand 'what makes this leader tick?'

In this chapter, I will first describe the manner in which a psychobiography is developed. I will then describe aspects of the personality study, emphasizing three political personality types. Two are quite common among political leaders: the compulsive personality, referred to above, and the narcissistic personality. The paranoid personality is much rarer but can be extremely dangerous when it occurs. Because of space limitations, I will only provide examples from profiles of how inferences can be drawn both from life course events and from behavioral observations.

More details on each element of constructing a Political Personality Profile can be found in the outline at the end of the chapter.

The psychobiography

The leader can be envisaged as residing within a series of fields, the cultural, historical, and political context of his country, the specific aspects of the leader's background, which shaped the individual, and the nature of the current political situation. The importance of that context cannot be overestimated. There is a profound difference in how personality will affect political behavior between a leader functioning in a collective leadership and a dictator in a closed system. The manner in which culture shapes expectations of the leader also shapes the formation and selection of the leader. The political leader who violates cultural norms will not long survive. In constructing a Political Personality Profile, the degree of constraint upon the political behavior of the leader by his role, the culture, and the nature of the political system is regularly examined.

The psychoanalytic framework of Erik Erikson (1963), which relates personality development to the cultural context, is extremely helpful as a model. It emphasizes the intimate dynamic relationship between the developing personality and the environment, and highlights the importance of the context in which the leader develops. Leader personality does not exist *in vacuo*; it is the leader *in context* that is our focus, both the context that shaped the leader's development and the contemporary context that continues to shape and influence behavior and decision-making. Thus, before even considering the particular circumstances surrounding the development of the future leader, one must understand thoroughly the culture, especially the political culture, in which the family was embedded.

The Political Personality Profile draws on the clinical case study methodology known as the anamnesis, which combines a psychobiography with a cross-sectional personality analysis. But the goal of the psychobiography developed to analyze political figures differs significantly from the analysis of the life course that psychiatrists develop to understand the traumatic events which predispose a patient to illness, for the goal is to understand how key life events have shaped the leader's personality, attitudes, and political behavior. Similarly, in the cross-sectional personality study of a political leader, the goal is not to specify dimensions of psychopathology, but rather to identify characteristic adaptive

styles and those aspects of cognition, attitudes, affect, and interpersonal relations which bear on specific elements of leadership functioning.

The manner of writing the psychobiography should prepare the reader for the detailed description of the political personality and analysis of leadership to follow. The psychobiography is a collapsing and expanding profile rather than a linear and chronological depiction of life events. It may be that one sentence captures years, while the details of a key afternoon may require several pages. The primary focus is on shaping events and experiences. Early leadership successes and failures are particularly important to identify and analyze in detail, as they are often endowed by the leader with exaggerated importance in guiding the leader's future political decisions.

Sources of identity

In the psychobiographic reconstruction, particular attention is given to specifying the sources of political identity. Erikson's emphasis on the formation and vicissitudes of personal identity is especially helpful in reconstructing the lives of political leaders, for as personal identity is consolidating, so too is political identity. This requires careful research into the preceding generations. For example, the influence of King Abdullah, the grandfather of King Hussein of Jordan, was profound. A charismatic man of towering political stature, Abdullah was ashamed of his son Talal, who suffered from chronic paranoid schizophrenia, so he started shaping his grandson to the role of future king. Young Hussein was at his grandfather's side on the steps of the Al Aqsa Mosque when Abdullah was struck down by an assassin's bullet. The 15-year-old boy too was struck by a bullet, but was reportedly saved from death by the medal on his chest that his grandfather had given him earlier that day – probably a powerful determinant of Hussein's sense of destiny.

Indira Gandhi recounted in her autobiography the influence of her grandfather Motilal Nehru, Congress party leader and prominent nationalist leader, and her father Jawaharlal Nehru, prime minister of India, who continued his father's struggle for Indian independence. When her parents were away in prison, as they often were during her politically tumultuous childhood, Gandhi indicated she did not play with dolls, but rather with toy metal soldiers. At the head of the column of soldiers was one with a white shield on which there was a red cross, suggesting her identification with Joan of Arc. She marched the clumsy soldiers into a fire again and again, suggesting the early foundation of her career long bent for conflict, and perhaps presaging her assassination by her

bodyguards. It is instructive to observe that she was characterized as 'the goddess of destruction' by her political opponents and was seen as a leader who regularly promoted political conflict, lacking her parents' conciliatory skills.

Key life transitions

Erikson follows the course of personality over the lifecycle, identifying the major crisis associated with each developmental epoch. Drawing on Erikson, Dan Levinson's work on the life course, *The Seasons of a Man's Life* (1978), is instructive in focusing on the three major life transitions: the Young Adult Transition, the Mid-Life Transition, and the Late Adult Transition. He emphasizes that the successful negotiation of each life transition requires successfully weathering the challenges of the previous life transition. I have developed the implications of Levinson's work for the influences of the lifecycle on the leader's political behavior (post 1980).

For example, Saddam Hussein's traumatic beginnings can be traced back to the womb. In the fourth month of his mother's pregnancy with him, his father died, probably of cancer. In the eighth month, his older brother died while undergoing surgery. His mother, understandably severely depressed, tried to abort herself of the pregnancy with Saddam and to commit suicide. She would not accept her newborn son in her arms, another sign of a grave depression. He was raised for the critical first years of life by his Uncle Khayrallah. When he was 3 years old, his mother remarried. Saddam went to her home, where his new stepfather abused him physically and psychologically. It is difficult to imagine more painful early years, which were the foundation of the wounded self, which underlay Saddam's grandiosity. I would submit that without understanding the magnitude of the traumas of Saddam's early years, it is simply impossible to understand the powerful forces within his political personality.

Childhood heroes and models are important to identify in the search for the foundation of political ambition, what Levinson has called the Dream: the crystallization of political ambition in adolescence that can serve as a lodestar. Young Anwar Sadat, for example, identified with Mohandas Gandhi and would cloak himself in a sheet, leading his goat around while on a self-imposed fast. I see this as the early foundation of his later role as peacemaker between Egypt and Israel that won him the Nobel Peace Prize.

The Dream may spur future greatness, but reactions to frustrated dreams of glory have led to intemperate acts that have been destabilizing as well. For instance, the Shah had written of his goal to transform Iran

into a modernizing Middle Eastern country. When he was informed in 1973 that he was ill with a slowly developing malignancy, he accelerated dramatically the pace of his efforts. Breaking with OPEC, he quadrupled the oil revenues pouring into the country's poorly developed infrastructure. This led to a revolution of rising expectations, which destabilized the social structure, leading to profound discontent, setting the stage for Khomeini's Islamic revolution. In his rush to accomplish his dreams before he died, he superimposed his personal timetable on the political timetable.

The role of the mentor is also extremely consequential. Young Iosif Dzhugashvili (who assumed the pseudonym Stalin 20 years later), oppressed by the rigors of the Orthodox seminary in Tbilisi, rebelled by smuggling in the works of Karl Marx and Vladimir Lenin. He came to idealize Lenin through his revolutionary writings and left the seminary to serve the cause of the revolution. But the contrast between Lenin as idealized model at a distance and the personal relationship to Lenin as mentor was striking. Initially a loyal protégé, increasingly Stalin became restive under Lenin's leadership, seeking power and authority for himself, leading up to a powerful confrontation between them when Stalin was in his early 40s, the height of the mid-life transition. Lenin subsequently suffered a disabling stroke, and Stalin went on to consolidate his power.

Psychologically salient issues

It is important to distinguish between those political behaviors deriving from the leader's role and those that engage his political personality. Discriminating which issues can be considered objectively and which strike deep psychological chords is crucial. For instance, President Chiang Ching Kuo was judicious and objective in his considerations of economic policy to create the economic miracle of Taiwan. His primary political mentor was his father, Generalissimo Chiang Kai-shek, ensuring that the issue of relationships with Mainland China could never be considered with the same rational objectivity. Progress towards ameliorating that conflicted relationship would have to await his death.

The leader who cannot adapt to external realities because he is rigidly adhering to an internally programmed life script has, in the terms of Harold Lasswell (1936), displaced his private needs upon the state and rationalized it in the public good. Inevitably the gap between the private needs and the public needs becomes the source of ineffective and/or conflicted leadership.

The political personality study

The goal of cross-sectional analysis is to identify and characterize the nature of the subject's political personality. The term 'personality' implies a *patterned relationship* among cognition, affect, and interpersonal relationships. Accordingly, the concept links belief systems, value systems, attitudes, leadership style, and other features. The nature of personality puts constraints upon the range of beliefs and attitudes, and the nature of relationships with the leadership circle, including who is chosen to serve as closest advisors, all of which influence political decision-making.

As with the longitudinal analysis in psychobiographic reconstruction, careful attention is given to all of the traditional elements considered in the clinical case study. These include the following: Appearance; Level of Activity; Speech and Language; Intelligence; Knowledge; Memory; Thought Content and Delusions; Drives and Affects (such as Anxiety, Aggression, Hostility, Sexuality, Activity and Passivity, Shame and Guilt, Depression); Evaluation of Reality; Judgment; Interpersonal Relations (such as capacity for Empathy; Identity and Ambivalence); and Characteristic Ego Defenses. Additional elements particular to political leadership are examined as well, including the following: Health (energy level, working hours, drinking, use of drugs); Cognitive/Intellectual Style, and the drives for power, achievement, and affiliation. The latter are important in attempting to identify whether the leaders sought their roles in order to wield power, to be recorded on the pages of history, or merely to occupy the seat of power with the attendant place in the limelight.

Ego defenses and personality types

It is particularly important to identify the characteristic pattern of ego defenses, for it is this repetitive manner of mediating between the subject's internal and external worlds that is at the heart of personality. The identification of patterns of ego defenses is not a matter of intuition but of recognition. Well-trained clinicians reliably identify the same characteristic ego defenses, but it does not require clinical training to be sensitive to and identify these patterns.

Each particular personality type has a characteristic array of ego defenses mediating between inner drives and the external world, and each has its own cognitive, affective, and interpersonal style. In evaluating ego defenses, it is useful to discriminate a hierarchy from primitive through mature. Vaillant (1992) has identified four levels of

defensive organization. The Psychotic Triad of Denial, Distortion, and Delusional Projection represent the most primitive level of psychological organization. The Immature defenses include the following: Projection, Passive Aggression, Acting Out, and Fantasy. The Neurotic (intermediate) defenses include Dissociation, Displacement, Isolation (or intellectualization), Repression, and Reaction Formation. Mature defenses include Suppression, Sublimation, and Altruism. Defenses do tend to aggregate, as exemplified by the so-called Psychotic Triad. This seriously disordered pattern is associated with paranoid psychoses and severe paranoid disorders. In contrast, the obsessive-compulsive personality pattern, which will be described in detail shortly, is associated with a much healthier array of Neurotic ego defenses.

Identifying a characteristic pattern of ego defenses is especially helpful in predicting behavior under stress, when these coping mechanisms can become exaggerated. This is particularly true in the face of serious illness and with increasing age. Thus the somewhat compulsive individual whose decision-making was unimpaired can become paralyzed by indecision, and the suspicious individual can become paranoid. Personality colors interpersonal relationships and thus can significantly distort relationships within the leadership circle. For instance, Beria was able to manipulate Stalin's paranoid tendencies to eliminate his own rivals. The fragile narcissist whose ego is intolerant of criticism may be impelled to surround himself with sycophants who can distort his appreciation of political reality.

In exaggerated form, each of these patterns can be psychologically disabling, at which time they would be considered as *personality disorders*. According to the standard psychiatric diagnostic reference, the Diagnostic and Statistical Manual of the American Psychiatric Association (DSM IV), the essential features of personality disorders are deeply ingrained, inflexible, maladaptive patterns of relating to, perceiving, and thinking about the environment and oneself that are of sufficient severity to cause either significant impairment in adaptive functioning or subjective distress. Thus they are pervasive personality traits and are exhibited in a wide range of important social and personal contexts. Since the stable pattern of defenses is also known as *character*, or the character armor (Reich 1933), personality disorders are also called character disorders.

Prominent examples of leaders with the full-blown disorders are found in the pages of history, particularly in closed societies led by dictators. Severe personality disorders are inconsistent with sustained political leadership in democracies, but, as noted above, under the stress of crisis

decision-making, personality patterns can temporarily show features of a disorder. Therefore, a number of the features in the following summary descriptions of the Narcissistic, Obsessive-Compulsive, and Paranoid personalities have obvious potential relevance to the decision-making and behavior of political leaders.

The narcissistic personality and its implications for leadership

It is probably not an exaggeration to state that if narcissistic characters were stripped from the ranks of public figures, those ranks would be significantly thinned. This label covers a broad range of behaviors. At the healthiest end of the spectrum are egotistical individuals with extreme self-confidence. But primitive narcissism, so-called malignant narcissism, represents an extremely severe and dangerous personality disorder. In addition to extreme self-absorption with an associated incapacity to empathize with others, it is characterized by a paranoid outlook, absence of conscience, and willingness to use whatever aggression is necessary to accomplish personal goals (Post 1993).

The essential features of the narcissistic personality disorder are a grandiose sense of self-importance or uniqueness. This tends to be manifested as extreme self-centeredness, egocentricity, and self-absorption. There is also a preoccupation with fantasies involving unrealistic goals, such as achieving unlimited power, wealth, brilliance, beauty, or fame, leading to an exhibitionistic need for constant attention and admiration and more concern with appearance than substance. These fantasies frequently substitute for realistic activity in pursuit of success. Even when the goals are satisfied, it is usually not enough; there is a driven quality to the ambitions. Abilities and achievements tend to be unrealistically overestimated, but minor setbacks can give a sense of special unworthiness.

The interpersonal relationships of narcissists are regularly and characteristically disturbed, vacillating between the extremes of over-idealization and devaluation. There is a constant need for reassurance and an exaggerated response to criticism or defeat. Because these individuals are so self-absorbed, they fail to empathize with others, who are seen as extensions of the self, there only to supply admiration and gratification. They regularly ignore the rights and needs of others. An individual is no longer perceived as psychologically useful can be dropped suddenly. Often extremely charming, the narcissist surrounds himself with admirers and requires a constant stream of adulation from them. They expect special treatment from others, expect others to do what they want, and will be angered when people fail to live up to their

unreasonable demands. There is accordingly a major inability to sustain loyal relationships over time.

A notable aspect of the narcissist in power is the manner in which this type of person seeks to gratify psychological needs through the exercise of leadership. Despite the apparent sustained devotion of their energies to socially productive endeavors, and 'selfless' rationales, the primary goal is actually to gain recognition, fame, and glory. This search for recognition and adulation springs from excessive self-absorption, intense ambition, and grandiose fantasies. But underlying and impelling this quest is an inner emptiness and uncertainty.

Kim Jong-il of the Democratic Peoples Republic of Korea demonstrates the impaired empathy of the narcissistic personality disorder. The average annual income of a North Korean is between \$900 and \$1000, and millions have starved in famines. Yet while Kim asks his people to sacrifice, he lives a remarkably hedonistic lifestyle in Pyongyang. According to the Hennessey Fine Spirits Corporation, in the 1990s, Kim annually spent between \$650,000 and \$800,000 on their most expensive cognac. Each grain of rice for 'Dear Leader' is inspected, and any with a minor defect is discarded. That he is insecure beneath his grandiose face is clear. Only about 5 feet 2 inches tall, he wears 4-inch lifts in his shoes and wears his hair in a pompadour to conceal his short stature.

The mirror image of the quest for adulation is sensitivity to slight and criticism. The narcissist is vulnerable and goes through complicated maneuvers to avoid being hurt. If the narcissist's self-concept of perfection and brilliance is to be sustained, no one can give him new knowledge, and no aspect of his understanding is to be faulted. Dogmatic certainty with no foundation of knowledge is a posture frequently struck. This profoundly inhibits acceptance of constructive criticism and leads to a tendency for the narcissistic leader to be surrounded by sycophants. The narcissist tends to devalue or even eliminate those who threaten his fragile self-esteem and, in subtle fashion, often plays one advisor off against another, to ensure that he is the major domo. This is particularly apt to stimulate the collective decision-making malady of 'groupthink.'

A vivid example is provided by Saddam Hussein, who demonstrated all of the characteristics of malignant narcissism. In 1982, when the war that he had initiated with Iran in 1980 was going very badly, he proffered a cease-fire to Iran's leader, Ayatollah Khomeini. But Khomeini said there would be no peace with Iraq until Saddam was no longer president. Saddam called a cabinet meeting and presented this dilemma. His sycophantic cabinet said in effect that 'you must stay on as president,

Saddam. Saddam is Iraq, Iraq is Saddam.' Saddam then went on to say, 'no, no, no, I want your frank, candid, and creative suggestions.' His Oxford educated minister of health Dr Ibrahim took him seriously and said, 'Well Saddam, you could withdraw from the presidency temporarily until our goal of peace is achieved, and then resume the presidency.' As the story goes, Saddam gravely thanked Ibrahim for his candor and arrested him on the spot. The minister's wife went to Saddam that night and pled with him, saying that her husband had always been loyal, and begged him to return her husband to her. The next morning he returned her husband to her in a black canvas body bag, chopped into pieces. This powerfully concentrated the attention of the remaining ministers who insisted that Saddam stay on as president of Iraq, and the war went on for another bloody 6 years.

The only central and stable belief of the narcissist is the centrality of the self. What is good for him is good for his country. It is hard to identify any consistent beliefs about the world because these tend to shift. Additionally, more than any other personality type, what the narcissist says should be viewed as calculated for effect. Accordingly, to place great weight on the analysis of core determining beliefs from speeches is apt to lead the unwary political analyst far astray, because words do not convey deeply held beliefs. Their only use is instrumental, to enhance personal position, to gain admiration and support. This attitude goes beyond 'naked' self-interest. The individual comes to believe that the national interest and national security are in fact crucially contingent upon his reelection or reappointment. For the narcissist, the problems are not threats to his country and what can be done to meet these threats but 'how can I use this situation to either preserve or enhance my own reputation?'

The obsessive-compulsive personality and its implications for leadership

The obsessive-compulsive (O-C) personality is frequently encountered in government and business executives, scientists and engineers, academic scholars and military leaders. Its strengths – organizational ability, attention to detail, emphasis on rational process – all can contribute to significant professional success. The O-C will have a sharp focus in examining the situation to get the facts but this preoccupation with details may show an inability to focus on 'the big picture.' Their everyday relationships tend to be serious, formal, and conventional, lacking charm, grace, spontaneity, and humor. There is an excessive devotion to work to the

exclusion of pleasure. Wilhelm Reich (1933) described these individuals as 'living machines.'

This personality places heavy reliance on the ego defense of intellectualization, emphasizing rationality and abhorring emotionality, which implies lack of control. Individuals with the O-C disorder tend to be preoccupied with matters of rules, order, organization, efficiency, and detail. This derives from an inordinate fear of making a mistake, for the goal of the O-C personality is to leave no room for error, to achieve perfection. Efficiency and perfection are idealized but, of course, never attained. Time is regularly poorly allocated, with the most important tasks left to the last moment. Decision-making is avoided, postponed, or protracted, with characteristic indecisiveness.

Usually such individuals are excessively conscientious, moralistic, scrupulous, and judgmental of self and others. Location in the interpersonal hierarchy is of great importance to individuals with this character type, who are preoccupied with their relative status in dominant-submissive relationships. Although oppositional when subjected to the will of others, they stubbornly insist on others submitting to their way of doing things and are unaware of the resentment their behavior induces in others. Relationships are serious and formal, demonstrating a restricted ability to express warm and tender emotions. Rigidity leads them to be described as dogmatic or opinionated. Such individuals are perceived as uninfluenceable.

The preoccupation with productivity and concentration imparts a special cast to the cognitive style and lifestyle of these individuals. They are immensely productive and show impressive abilities to concentrate on their work, often cranking out huge volumes of work, especially in technical areas. But everything seems laborious, determined, tense, and deliberate; there is a quality of effortfulness, leading to the frequent characterization of them as 'driven.' But the driver for the O-C is his own harsh taskmaster. He regularly tells himself (and others) what he 'should' do; the language of 'want' is alien. While these directives are burdensome, they also provide clear guidelines for behavior. The guarded state of attention, the inability to relax, the preoccupation with 'should' are all in the service of avoiding losing control. There is a tight lid on feelings, an avoidance of impulse or whim.

This has major consequences for decision-making. The preoccupation with doing what is right places a premium on avoiding mistakes. Yet the O-C personality tends to see a world in shades of gray, characterized by complexity and subtlety. Accordingly, O-Cs often have difficulty in making clear choices and reaching closure, searching for additional

evidence to ensure that they are not making a mistake. They follow the imperatives to 'act only after gathering as much information as possible' and 'preserve one's options as long as possible,' preferring procrastination to the dangers of hasty action or premature closure. This agony can be forestalled if there is a rule that can be applied. If there is no formula, however, the O-C will become quite anxious. New and unanticipated situations are particularly threatening.

The O-C will want to receive raw data, to see the minutiae about almost everything, and to become immersed in as many details as possible in a quixotic quest to somehow fully understand the issue. They have a great deal of difficulty delegating and relying upon subordinates, who, after all, might make a mistake. Crisis decision-making, where there is uncertainty or ambiguity, is especially difficult. Eventually, the O-C becomes overwhelmed and begins to think of issues in terms of data.

During an impasse in the Camp David negotiations, President Jimmy Carter made effective use of the profiles of Menachem Begin and Anwar Sadat that my unit at the CIA had prepared for him. Begin, who had been characterized as a compulsive tending to become preoccupied with details, was being intractable and refusing to compromise. Carter brilliantly exploited his understanding of Begin as an O-C personality, stating to him, 'Your excellency, President Sadat is concerned that we will become so bogged down in details that we'll lose sight of the big picture.' As reported by Carter, Begin drew himself up stiffly and responded, 'I too can focus on the big picture. We'll leave the details to our subordinates.' And they got past this impasse.

Dominated by a strong conscience, the O-C personality is a man of his word, like the Latin American leader I characterized at the start of this chapter. When he has made a commitment in negotiations, he can be relied upon, in contrast to the Narcissistic Personality, who can reverse himself as circumstances dictate. Moreover, to the extent that he has committed to writing his policy goals and preferences, these can be taken as a reliable map of intentions.

The paranoid personality and its implications for leadership

The essential features of the paranoid personality disorder are a pervasive and long-standing suspiciousness and mistrust of people in general. Individuals with this disorder are hypersensitive and easily slighted. They continually scan the environment for clues that validate their original suspicions and dismiss evidence that disconfirms their fearful views. A striking quality is pervasive rigidity. The suspicious person searches repetitively only for confirmation of danger, whereas

the psychologically healthy individual can abandon suspicions when presented with convincing contradictory evidence. Attempts to reassure the paranoid or reason with him will usually provoke anger and the 'helpful one' may become the object of suspicions as well. This is because the paranoid position is restitutive, that is, designed to compensate for feelings of loss or insignificance. It is preferable to be the very center of a conspiracy than to be alone and insignificant. There is a world of hidden motives and special meanings. They have a readiness to counterattack against a perceived threat, and can become excited over small matters, 'making mountains out of molehills.'

The paranoid is hyper-vigilant, ever alert to a hostile interpersonal environment and ready to retaliate. Always expecting plots and betrayal, his antennae constantly sweep the horizon for signs of threat, often generating fear and uneasiness in others. This requirement for enemies explains why paranoia is the most political of mental disorders. The only defense in such a dangerous world is to rely on no one, leading to an exaggerated emphasis on independence and autonomy. In a new situation, paranoids intensely and narrowly search for confirmation of their bias with a loss of appreciation of the total context. They usually find what they anticipated finding. Insofar as the paranoid intentionally seeks out only data which confirms his premise of external danger, and systematically excludes evidence to the contrary, his evaluation of reality is often skewed by the ego defense of Projection – the attribution to external figures of internal motivation, drives, or other feelings that are intolerable and hence repudiated in oneself.

Paranoids tend to be rigid and unwilling to compromise. Priding themselves on always being objective, unemotional, and rational, they are uncomfortable with passive, soft, sentimental, and tender feelings. They avoid intimacy except with those they absolutely trust, a minute population. The paranoid guards against losing control of his feelings. Being on guard at all times blocks spontaneity, and the absence of spontaneity clearly inhibits creative expression. There can be no humor or playfulness. Keenly aware of rank and power, superiority or inferiority, they are often jealous of and rivalrous with people in power. They avoid participating in a group setting unless they are in a dominant position. There can be no yielding to pressure or authority. The exaggerated need for autonomy has significant implications for leadership style but also affects cognitive style.

Thus the paranoid is simultaneously defending himself against external danger and against internal impulses, a burdensome and exhausting psychological war on two fronts. As internal tension builds,

suspiciousness grows, and through the process of projection a *more manageable* external threat is constructed. The internal persecutor is infinite in quantity. The external enemy is finite and can be destroyed. The individual then has a state of heightened alertness, a state of continuous guardedness against the *now external* danger.

Of all the personality types, this is the one most motivated to maintain internal consistency among cognitive beliefs, often at the expense of an 'objective' examination of new information. The paranoid typically holds a very strong, rigidly entrenched belief system with a vivid and central image of the adversary. As one might suspect, the adversary is seen as inherently and pervasively evil and an incorrigible threat to one's own personal/national interest. The paranoid personality, by definition, sees the world in polarized terms, as a Manichean universe divided into two camps – allies and adversaries. Neutrals are impossible.

There is a powerful tendency to exaggerate greatly not only the hostile nature of the adversary's intentions but also the adversary's political and military skill. The paranoid personality tends to view the adversary as highly rational, highly unified, in total control of all his actions. People or nations are never compelled to do things by virtue of circumstances or by accident. War would never emerge in a crisis for inadvertent reasons. Actions are always a product of the adversary's negative qualities; war occurs because of nefarious, aggressive motivations. There is no such thing as a defensive action by the adversaries – all their actions are necessarily aggressive. One can never safely assume that the adversary's military potential is so small that it will never become a threat, even if it is not one now. The world is a conflictual place and the source of conflict is the evil nature or character of other nations or people.

Precisely because of the rigidity of these beliefs and the central importance of this image of the adversary, the paranoid's worldview is heavily biased in favor of worst-case analysis. Faced with the need to make a decision, the paranoid personality will manifest a strong tendency to act sooner rather than to procrastinate and has a strong preference for the use of force over persuasion. The paranoid may even initiate a crisis or a war out of the belief that preventive action against the adversary is necessary. The point is to alter leadership or capabilities rather than try to alter behavior. In fact, the information search pattern will be exclusively tactical in nature because the long-term objectives of the adversary are already known. A related topic of interest will be information relating to the 'enemy within' or 'fifth column activity.' The adversary is believed to be very creative and devious in this sort of covert subversion, and people

of one's own country who do not fully share the views of the paranoid leader are believed to be either suspect themselves or, at best, naïve.

The paranoid will tend to adopt a management style that rests on the assumption that one cannot trust any one source of information or any one concentration of power. So to garner diverse information and, most importantly, to prevent the rise of any potential internal threats, the paranoid leader plays one advisor or one bureaucracy off against another one. They will typically not be satisfied with the analyses and conclusions of people working under them. The manipulative subordinate can take advantage of the paranoid leader's suspiciousness to plant suspicions concerning bureaucratic rivals, as did Beria with Stalin.

Conclusion

Case studies of individual leaders provide a particularly valuable longitudinal perspective that offers a framework for understanding the manner in which previous life experiences influence political behavior and for distinguishing between political behaviors, which are role dependent and those which reflect strong personality influences. The Political Personality Profile identifies enduring aspects of leader personality, including cognitive, affective, and interpersonal elements. A key aspect is the combination of psychobiographic and psychodynamic approaches in order to understand themes ingrained during adolescence, which continue to influence throughout the lifecycle. Dreams die hard, and pursuit of the dreams of glory formed during adolescence can drive a leader throughout his lifetime, having special force at the mid-life transition and during the later-years transition.

Because personality is stable over time, the longitudinal approach helps identify enduring patterns. Three important leader personality types – the Narcissist, the Obsessive-Compulsive, and the Paranoid – and their implications for political behavior have been described in order to illustrate this principle. Because these personality patterns are so deeply ingrained, they can be detected early in a political career, and can reliably be predicted to continue to affect leadership behavior throughout the political career, and become intensified with stress. When present, they permeate all aspects of political behavior, including crisis decision-making, strategic decision-making, negotiating behavior, worldview, and relationships with the leadership circle.

Having described these character types in detail, it is important to emphasize that most individuals, and most leaders, possess a broad array of characteristics that do not fit one pure type. Rather, it is the

predominance of one style over another that affects outcomes. The healthy leader has personality characteristics that contribute to effective leadership, to sound decision-making, to accurately diagnosing of the environment, and to working effectively with a circle of advisors chosen for their expertise and wisdom, from whom the self-confident leader can learn and take wise counsel.

APPENDIX: FRAMEWORK FOR AN INTEGRATED POLITICAL PERSONALITY PROFILE

Part I: Psychobiography

1. Cultural and historical background – describe constraints of the political culture on the role of leader.
2. Family origins and early years
 - a. Family constellation: grandparents, parents, siblings, relationships, politics of family
 - b. Heroes and models.
3. Education and socialization
 - a. Climate in country
 - b. Student years, leadership.
4. Professional career
 - a. Mentors
 - b. Early career
 - c. Successes and failures.
5. The subject as leader
 - a. Key events
 - b. Crises
 - c. Key political relationships, influences.
6. Family and friends.

Part II: Personality study

1. General Personal Description
 - a. Appearance and personal characteristics (lifestyle, work/personal balance, working hours, hobbies, recreation, etc.)
 - b. Health (include energy level, drinking, drug use).

2. Intellectual capacity and style
 - a. Intelligence
 - b. Judgment
 - c. Knowledge
 - d. Cognitive complexity.
3. Emotional reactions
 - a. Moods, mood variability
 - b. Impulses and impulse control.
4. Drives and Character Structure
 - a. Identify Personality Type (if possible)
 - b. Psychodynamics
 - i. Self-concept/self-esteem
 - ii. Basic identification
 - iii. Neurotic conflicts
 - c. Reality (sense of/testing/adaptation to)
 - d. Ego defense mechanisms
 - e. Conscience and scruples
 - f. Psychological drives, needs, motives (discriminate to degree possible among drive for power, drive for achievement, drive for affiliation)
 - g. Motivation for seeking leadership role (to wield power, to occupy seat of power, to achieve place in history).
5. Interpersonal Relationships
 - a. Identify key relationships and characterize nature of relationships
 - i. Inner circle, including unofficial advisors, 'kitchen cabinet'
 - ii. Superiors
 - iii. Political subordinates
 - iv. Political allies, domestic and international
 - v. Political rivalries, international adversaries.

Part III: Worldview

1. Perceptions of political reality (include cultural influences/biases)
2. Core beliefs (include concept of leadership, power)
3. Political philosophy, ideology, goals, and policy views (domestic, foreign, and economic policy views; include discussion of which

issues most interest the leader, in which issue areas his experience lies, and which issues are particularly salient for his political psychology). Note that not all leaders have a core political philosophy or body of governing political ideas.

4. Nationalism and identification with country.

Part IV: Leadership style

1. General characteristics (include discussion of the role expectations, both general public and elite, placed on the individual emphasizing their political and cultural determinants and leader's skill in fulfilling them)
 - a. How does subject define his role?
 - b. Relationship with public
 - c. Oratorical skill and rhetoric.
2. Strategy and tactics-goal directed behavior
3. Decision-making and decision-implementation style
 - a. Strategic decision-making
 - b. Crisis decision-making
 - c. How does he use his staff/inner circle? Does he vet decisions or use them only for information? How collegial? Does he surround himself with sycophants or choose strong self-confident subordinates?
 - d. Dealing with formal and informal negotiating style.

Part V: Outlook

1. Note particularly political behavior closely related to personality issues. Relate personality to key issues emphasizing in which direction the psychological factors point. Estimate drives, values, and characteristics that are the most influential.
2. Attempt to predict how the individual will interact with other political figures, including opposition leaders and other key foreign leaders.

10

Content Analysis

Margaret G. Hermann

Only movie stars, hit rock groups, and athletes leave more traces of their behavior in the public arena than politicians. Few of a US president's or a British prime minister's movements or statements, for example, escape the media's and archivists' notice. With 24/7 coverage and the Internet, what leaders from around the world discuss is often beamed into our televisions and put onto the Web. Through content analysis, such materials help us learn more about essentially unavailable public figures, because it does not require their cooperation. Computer-assisted software (such as Atlas.ti, Nudist, Profiler+) and the increase in Internet sources that record material from news services, television, elites' papers, and archives have improved the ease and reduced the time necessary for conducting such analysis.

Content analysis involves developing 'a set of procedures to make inferences from text' (Weber 1990: 19); it is a method 'capable of throwing light on the ways [people]... use or manipulate symbols and invest communication with meaning' (Moyser and Wagstaffe 1987: 20). A series of eight steps and a similar number of decisions need to be considered before beginning the analysis. (1) Does your research question involve extracting meaning from communications? (2) What kinds of materials are available and how accessible are they? (3) Does what you are interested in studying lend itself more to a qualitative or quantitative analysis? (4) Do you view the materials as representational or instrumental in understanding the subjects you are studying? (5) What is your unit of analysis, and what coding rules and procedures do you plan to use? (6) Can one contextualize to take into account situation, culture, and history? (7) Can others replicate your analysis? (8) Does the analysis capture what you are interested in learning about? To illustrate

these types of decisions, I will show how content analysis can be used to determine leaders' responsiveness to political constraints.

Step 1: considering the research question

Does your research question involve extracting meaning from communication? I want to find out how much leaders discuss exerting control and influence over their environment and the constraints that the environment poses (as opposed to focusing on the need to be adaptable to the situation and the demands of various constituencies). For example, research shows that leaders who talk about challenging constraints are more intent on meeting a situation head-on, achieving quick resolution to an issue, being decisive, and dealing forcefully with the problem of the moment (Driver 1977; Hermann 1984; Tetlock 1991; Suedfeld 1992; Keller 2005). In effect, how political leaders talk about the constraints in their environments helps to shape their expectations, strategies, advisory systems, and the actions they urge on their governments (or other types of political units). Thus the research question involves extracting meaning from the communications of public figures. My response to the first question in contemplating doing a content analysis is 'yes.'

Step 2: selecting material

Consider the wide range of materials that are available for content analysis: books, films, pamphlets, party manifestoes, television programs, speeches, interviews, children's readers, newspapers, election commercials, blogs, diaries, letters, open-ended interviews, survey responses, cartoons. Anything that is intended to communicate a message is usable as material for content analysis. Moreover, the material does not need to involve words. Content analysis can also be used to examine nonverbal behavior (Hermann 1979; Hermann and Hermann 1990). The type(s) of relevant material generally depends on the research question.

Because of my interest in ascertaining how political leaders – in particular, heads of governments from around the world – interpret the constraints they perceive in their environments, I have focused my content analysis on their speeches, press conferences, and interviews. I have found the need to exercise some caution in using speeches since such materials often are written by speechwriters or staff members. In certain types of political systems, speeches may even be crafted by

committee, reflecting the nature of the consensus that was achievable at that time.

Speeches are usually designed for particular audiences and occasions. Care and thought have generally gone into what is said and how it is said. During election periods, for example, the viewer can almost see candidates turning on their mental tape recorders as they move from place to place. In speeches, the leaders are presenting their public persona – how they would like to be perceived by the public. They are exhibiting what they believe will win votes, mobilize support for particular positions, and improve or maintain high approval ratings. Research suggests that such public statements reflect what the leader wants and is pledging to be (Winter and Stewart 1977; Winter *et al.* 1991; Winter 2005). This same research also suggests that there is a linkage between the motives expressed in speeches and how presidents engage in leadership during their time in office.

Press conferences and interviews, in contrast to speeches, are generally more spontaneous. Although there is often some prior preparation (such as consideration of what questions might be asked and, if asked, how they should be answered), leaders are on their own. During the give and take of a question and answer period, in particular, leaders must respond quickly without props or aid. What they are ‘really’ like can influence the nature of the response and how it is worded. The interview lets the analyst come closer to learning about the private persona of the leader – what the leader is like behind the scenes in the decision-making process. Of course, the most spontaneous interviews are those when the leader is caught in an impromptu setting, such as when leaving a meeting. Because I am interested in assessing leadership style and in learning as much as possible about the more personal side of the leader, interviews and the give-and-take in press conferences are my material of preference. (For research that explores the differences between speeches and interviews in the assessment of public figures, see Hermann 1986; Winter *et al.* 1991; Schafer 2000.)

Interviews, press conferences, and speeches of leaders are available in a wide variety of sources. Such materials for political figures located in governments outside the United States are collected in the *Foreign Broadcast Information Service Daily Report* and are reported by other governments’ information agencies on their websites. Interviews with political elites who reside within the United States are often found in such newspapers as the *New York Times* and *Washington Post* as well as in weekly news magazines and on the websites for weekly television news programs. *Lexis-Nexus* provides such materials from a diverse number of

news sources. Press conferences and other interviews with US presidents as well as their speeches can be found in each one's *Presidential Papers*.

Because I am interested in learning about what a political leader is like, it is particularly important in collecting speeches and interviews that I locate verbatim materials – the full text as spoken by the leader. At times, newspapers and magazines will overview or edit interviews with leaders – and even speeches – making it difficult to know how representative the material reported is. I am less interested in what the particular media outlet believes will sell newspapers or magazines than in how the leaders presented themselves. (Of course, other researchers may find that question worth pursuing, and checking what is left out and included by the media can turn into a research question amenable to content analysis.)

In the course of determining how responsive close to 200 national political leaders are to the constraints in their environments, it has become evident that the analyst can develop an adequate assessment based on 100 interview responses of 150 words or more. Any profile will suffer if it is based on less than 50 responses. (Confidence in one's profile, of course, increases with more interview responses.) To insure that responsiveness to political constraints is not context-specific, the interview responses that are analyzed should span the leader's tenure in office as well as have occurred in different types of settings and focus on a variety of topics. Collecting and categorizing interview responses by time, audience, and topic provides a means for assessing how stable what you are discovering is. Such data help the analyst know how relatively sensitive or insensitive to the context a particular leader is. With the advent of computer software (noted above) that can assist in the content analysis, it is feasible now to collect as wide a variety of machine-readable material as is available and to worry less about how much material is enough.

One question that is always raised about the use of speeches, press conference, and interviews in the study of leaders and leadership, particularly in countries outside the English-speaking world, concerns the effects of translation. Words mean different things in different languages and cultures. How can we assume that what I am considering an indication of the leader challenging constraints, for example, means the same in the United States and Uganda? One way of dealing with this issue has been to use material from such sources as the *Foreign Broadcast Information Service Daily Report* that involves native-born translators on site in the various countries whose media is being monitored. When these individuals do not know how to translate a word into English, they bracket

the original word so that the reader knows the translator is unsure what the equivalent English word is.

Those of us interested in studying the characteristics of political leaders through content analysis have also run some tests having persons we have trained from other countries to do the content analysis in their original language while we do the same on the English translation. So far such tests have been conducted with documents in French, Swahili, Russian, and Mandarin Chinese. On average, the overlap in the results of the tests has been 87 per cent. Where there were differences, we have worked on considering why they were occurring, and where possible have developed rules to help minimize the so-called mistakes or added more words to our dictionaries in the case of the computer software.

Steps 3 and 4: deciding on the nature of the content analysis

After selecting the material to be analyzed, there are several decisions the researcher must make regarding what kind of content analysis to do. Does what you are interested in studying lend itself more to a qualitative or quantitative analysis? What kinds of assumptions are you making about the representational or instrumental nature of the material? The answers to these two questions are often interrelated, as I will illustrate.

Research about the degree to which leaders believe that they can influence what happens (or have a need for power) suggests whether they will challenge or respect the constraints that they perceive. Note that the focus is on 'the degree to which' leaders have 'more or less' of certain characteristics that would explain whether they are respecters or challengers of constraints. Therefore, doing a more quantitative analysis is appropriate. An assumption is made that the more frequent use by leaders of certain words and phrases in their interview responses indicates their belief that they can influence what happens (or have a need for power), the more salient such content is to them, and the more representative it is of what they are like. This answers both questions raised above. The analysis is going to be more quantitative in nature and representational in form. I presume that leaders' 'inner feelings are accurately reflected in what they say' (Holsti 1969: 32) and that frequency represents salience to the speaker.

Compare these responses to someone involved in propaganda analysis – say, examining the statements of a person like Osama bin Laden – and trying to ascertain when a particular action is likely to take place. Here the focus of attention is likely to be the 'presence or absence' of particular themes and targets. The analysis will be more qualitative in

nature with an assumption being made that the interview responses or speeches being studied are instrumental in nature – that one cannot take the message at its face value but needs to examine ‘what it conveys given its context and circumstances’ (Holsti 1969: 32). In such an analysis, what is not said may be as important as what is said, particularly if a theme emphasized over a length of time all of a sudden disappears. Moreover, sometimes in this type of analysis, the conjunction of several themes or targets may become significant, such as when suddenly two ideas appear together that have been separated before.

Thus content analyses can range from being more qualitative to more quantitative, depending on whether the focus of attention is on the presence or absence of certain characteristics or on the degree to which the speaker exhibits the characteristics. It is, of course, possible to use numbers when a phenomenon is present versus absent but it is harder to judge the degree without the use of some quantification.

Most studies that use content analysis to assess what political leaders are like are representational in nature, because they are interested in inferring from what the leaders say something about what they are like, be it leadership style, beliefs, or motives. But the difference between considering speeches and interviews of leaders as representational versus instrumental reflects the ongoing consideration among those involved in such analysis of the distinction between public and private personas. After all, some might argue that everything a political leader says is instrumental – focused on persuading others. As Harry Truman is supposed to have said, ‘I spend my day trying to persuade people to do things they should be doing anyway.’ It is for this reason that I insist on examining interview responses that differ in degree of spontaneity and even in comparing the analysis of speeches to those of interviews. We know that speeches are often more planned and, thus, probably more reflective of what the leader wants others to perceive. The results from speeches can be used as a baseline against which to compare what one learns from analyzing the most spontaneous interviews, which are probably more representative of the person *qua* person. Analyses of things like Nixon’s Watergate Tapes and Dag Hammarskjöld’s *Markings* (his poems written at the end of the day) suggest that the more spontaneous interview responses come close to matching the persona that is evident outside the public’s view (see Schafer 2000).

Generally in the assessment of what political leaders are like, the focus is on doing a frequency analysis. The assumption is made that the more frequently leaders use certain words and phrases in their interview responses, the more salient such content is to them. There are, however,

other types of content analysis that one can use, namely contingency analysis and evaluative assertion analysis.

Contingency analysis has as its focus exploring when words, phrases, or themes appear together – when one is contingent on the other. For example, in examining what types of reactions the political leaders involved in a transit strike were likely to exhibit when their stress levels were high, the author (Hermann 1977) studied the transcripts of interviews with these leaders for when indicators of stress were paired with defensive reactions suggesting denial or confrontation. Interest centered on the conjunction of the appearance of stress and what the leader did.

Evaluative assertion analysis focuses on the strength with which something is said. When a political leader talks about his administration, for instance, what kinds of adjectives does he use? His is a strong, coherent, effective administration or a problem-ridden, veto-prone, distracted administration. The adjectives can be scored according to the degree to which they are positive or negative, suggest cooperation versus conflict, indicate strength versus weakness. This type of analysis was used in the study of the leaders involved in the July 1914 crisis that led to the outbreak of World War I (Holsti 1972). Of interest was what the governments did when the leaders' perceptions of the hostility of the other parties increased.

Step 5: determining the unit of analysis and coding

Okay, decisions have been made regarding which way one wants to go on the qualitative–quantitative continuum, with reference to assumptions dealing with how representative versus instrumental the analysis will be, and about the type of content analysis. At issue next are the units of analysis one is going to use. What exactly are you going to code? In my case, I know the material will be a speech or interview, but what is it about the speech or interview that I plan to examine?

Units can range from words to phrases, sentences, paragraphs, themes, and whole documents. With an interest in discovering which political leaders challenge constraints as opposed to respect them, I want to learn when leaders take responsibility for an action – when they believe that they have some control over what happens and have a need to have influence. Thus, in examining interview responses, my focus is on verbs or action words. Assuming when leaders take responsibility for planning or initiating an action that they believe they have some control over what happens, of interest is actions proposed or taken by the leader or by a group with which the leader identifies. The content analysis

therefore is centered on verbs whose subject is the leader or such a *group*. Do such verbs *suggest* that the leader is *planning*, *initiating*, or *taking* responsibility for an action? Or do they *suggest* that the leader is *reacting* to another's initiative, *stating* a fact or opinion, or *commenting* on what is happening? I am interested in the percentage of time the verbs in an interview response indicate that the speaker or *group* with whom the speaker identifies has taken responsibility for *planning*, *initiating*, or *taking* an action.

Consider some other ways of studying political leaders. David Winter (1992, 1995) wants to assess leaders' motivation and is interested in the stories that leaders tell in their speeches and in the kinds of themes represented in those stories. He codes for themes having to do with achievement, power, and relationships. Peter Suedfeld (1992; Suedfeld and Wallace 1995; also Young and Schafer 1998) studies how integratively complex political leaders are. He observes how leaders *integrate* ideas in the whole speech. Are they capable of *integrating* discrepant ideas in their speeches or do they tend to stop at *differentiating* among thoughts and objects? Suedfeld has developed a seven-point scale to rate speeches on the degree of integration of ideas that is present in what is discussed.

Note that in each of these examples, the researchers have done something to control for the length of the material; they have chosen a system of enumeration or unit of *aggregation*. Since speeches differ in the number of words they contain, and interviews in the number of questions asked, comparison becomes easier when these differences are taken into account. I use the percentage of time that an action verb meets the specified criteria among the number of such verbs in an interview response. That is, when the leader could have indicated a belief in being able to control events, how often did he or she? Winter determines the number of themes he finds in speeches per 1000 words. Suedfeld attaches a particular score on his seven-point integrative complexity scale to each speech.

Once the decisions are made regarding the units of analysis and enumeration, researchers are faced with *designing* the rules and procedures they are *going* to follow in drawing inferences from the materials. Let us elaborate on the example of *differentiating* between leaders who *challenge* and *respect* constraints.

As mentioned above, the belief that one can control what happens and a need for power seem to differentiate political leaders who *challenge* constraints from those who do not (McClelland 1975; Winter and Stewart 1977; Hermann 1980a; Walker 1983; Hermann and Preston

1998; Kaarbo and Hermann 1998; Keller 2005). Political leaders who are high in their belief that they can control what happens and in their need for power have been found to challenge the constraints in their environments; indeed, they generally try to push the limits of what is possible. Moreover, they are skillful both directly and behind the scenes in getting what they want. Leaders, however, who are low in these two traits appear to respect, or at the least accede to, the constraints they perceive in their environments and to work within such parameters toward their goals. Leaders who are moderate on both these traits have the ability of moving either toward challenging or toward respecting constraints, depending on the nature of the situation; they will be driven by their other characteristics, and what they believe is called for by the context.

What about a leader who is high on one trait and low to moderate on the other? Leaders who are high in the belief that they can control events but low in need for power are found to take charge of what happens and to challenge constraints, but they do not do as well in reading how to manipulate the people and setting behind the scenes to have the desired influence. Such leaders are not as successful in having an impact as those high in both the traits. They are too direct and open in their use of power, signaling others on how to react without really meaning to do so. And what about the leaders who are low in belief that they can control events but high in need for power? These individuals also challenge constraints but they feel more comfortable doing so behind the scenes, in an indirect fashion rather than out in the open. Such leaders are especially good in settings where they are the 'power behind the throne,' where they can pull the strings but are less accountable for the result.

In coding for the belief that one has some control over events, as was mentioned earlier, the focus is on verbs or action words. We assume that when leaders take responsibility for planning or initiating an action, they believe that they have some control over what happens. Consider the following response from a leader in an interview: 'I am sending our troops to the border to quell the uprising.' Compare this response to another: 'We are restrained by their actions.' In both cases, the leader or a group with whom he or she identifies is the actor, but in the first case the leader is initiating an action while in the second the leader is remarking about not being able to act. The first would be coded 'one' as indicative of a belief in being able to control what happens; the second does not reflect such a belief and would be coded 'zero.'

A score on this trait is determined by calculating the percentage of time the verbs in an interview response indicate that the speaker (or a group with whom the speaker identifies) has taken responsibility for planning

or initiating an action. The overall score for any leader is the average of these percentages across the total number of interview responses being examined. This way of coding for belief in one's ability to control what happens was developed by examining the 'locus of control' measure constructed by Julian Rotter (1966; see also Lefcourt 1976) and is used extensively in the study of leadership in a variety of contexts (for a review, see Nahavandi 2003). Specifically, this set of coding rules emulates the items used by Rotter to assess an individual's internal locus of control that focused on taking the initiative and engaging in action. The items that evidenced a low internal locus of control were centered on the effects of fate, luck, chance, and the presence of constraints.

Need for power indicates a concern for establishing, maintaining, or restoring one's power or, in other words, the desire to control, influence, or have an impact on other persons or groups (Winter 1973). As with the previous trait, my coding focuses on verbs. Is the speaker proposing action that attempts to establish, maintain, or restore his or her power? Some of the conditions where need for power is scored are when the speaker (1) proposes or engages in a strong, forceful action such as an assault or attack, a verbal threat, an accusation, or a reprimand; (2) gives advice or assistance when it is not solicited; (3) attempts to regulate the behavior of another person or group; (4) tries to persuade, bribe, or argue with someone else so long as the concern is not to reach agreement or to avoid disagreement; (5) endeavors to impress or gain fame with an action; or (6) is concerned with his or her reputation or position. Once again the focus is on actions proposed or taken by the leader or a group with whom he or she identifies. A score on need for power is determined by calculating the percentage of time the verbs in an interview response indicate that the speaker or a group with whom the speaker identifies has engaged in one of these six behaviors. The overall score for any leader is the average of these percentages across the total number of interview responses coded.

Step 6: contextualizing the information

Having decided on a unit of analysis and developed a set of coding rules and procedures, it is time to consider how to contextualize the information you are collecting to account for the nuances and complexities that are part of any political phenomenon. Currently, I have information on close to 200 political leaders around the world (starting from 1945) including members of cabinets, revolutionary leaders, legislative leaders, leaders of opposition parties, and terrorist leaders in addition to heads

of governments. Once a leader's interview responses have been coded and overall scores have been calculated across the interview responses, it is time for me to put the scores into perspective by comparing them with those of other leaders. Without doing such a comparison, there is little basis on which to judge if the particular leader's traits are unusually high or low or about average.

The issue is deciding what group of leaders to use as the comparison. A so-called norming group defines what is more usual – the norm – among a particular group of leaders (as opposed to more extreme or different). I am interested in whether or not a leader's scores are one standard deviation above or below the mean for the norming group, indicative of being higher or lower than the average leader in the comparison group. It is possible to break down these sets of leaders into particular regional, country, and cultural groups and compare the leader under study to a group that is close geographically and culturally.

For example, consider the scores on 'belief that one can control events' and 'need for power' for three Iranian leaders – the Ayatollah Khamenei, former President Khatami, and current President Ahmadinejad – when compared to a group of Middle Eastern leaders and one containing other Iranian leaders (Hermann 1999). Khamenei's scores (low in belief that he can control what happens and high in need for power) suggest that he will challenge constraints but do so in an indirect, behind the scenes manner. And, indeed, although Khamenei does have ultimate authority in the Iranian political system, he prefers to maintain control and maneuverability by not being 'out in front.' The former President Khatami's scores (low in belief that he can control events and low need for power) indicate a focus on respecting constraints and working within the system for change. In contrast, the current President Ahmadinejad's scores (high in belief that he can control what happens and low need for power) suggest that he challenges constraints but does so in an open and direct manner, often signaling where he is going and allowing others behind the scenes to out-manuever him when it comes to actually wielding influence. Even though they are (or were in the case of Khatami) in positions of power in the same country, these leaders show different ways of responding to constraints that have implications for how policy will be made.

Research also suggests that while some leaders use contextual cues to determine what they do and, thus, may evidence changes in their trait scores depending on the nature of the situation, other leaders' styles are fairly stable across situations. By examining diverse material on a political leader, it is possible to determine how stable his or her leadership

traits are. If there is variability in the scores, then we can determine if the differences give us insights into how political leaders adapt to the situations in which they find themselves. Selecting materials that cover a leader's tenure in office, and even before if such material is available, enables us to contextualize the profile of a leader. Computer software comes in handy, as it facilitates dividing the material in different ways and exploring how such differences affect leaders' characteristics. (In addition to Profiler + described in Young 2000, check out the Atlas.ti and Nudist programs described in Barry 1998.)

Another research question is whether leaders exhibit different styles when topics change; maybe they perceive themselves to be experts in some domains and not in others. Consider Jiang Zemin of the People's Republic of China. His scores regarding belief that one can control events differ by topic affecting, in turn, his willingness to challenge constraints. In considerations of regional politics, Jiang had a greater sense that he could control what happened than with regard to China's relations with the United States. He was willing to challenge constraints openly when the focus was on Chinese influence in the region but more likely to respect constraints when it came to dealings with the United States (Hermann 2000).

In order to examine if and how a leader's traits may differ by the substance being discussed, it is necessary to ascertain what topics are covered in the material under analysis. What the leader is talking about in each interview response that is being coded? It is generally possible to arrive at a set of categories by checking where the topics covered are similar and which topics are discussed the most. Topics that are covered only sporadically in the material are good candidates for combining into more generic categories. For instance, technological development might be collapsed into a category called economic issues. (Note the similarities with the discursive analysis in both Neumann and Dunn in this book.)

Interviews with political leaders are done in a variety of settings and, thus, are often targeted toward different audiences. Leaders often challenge constraints directly when facing their domestic public while challenging constraints more indirectly and behind the scenes when targeting leaders in other countries. Slobodan Miletovic is an interesting case (Hermann 1999). His scores suggest that as President of Serbia he generally challenged constraints domestically and regionally in an indirect manner, but when his focus was on the United States or Britain, he became more respecting of constraints and highly interested

in interacting with the representatives of these countries to ascertain what would 'sell.'

To examine the effects of audience on a leader's scores, it becomes important to note who the interviewer is and where the interview is taking place. For example, in profiling a head of government, a researcher will want to record if the interview involves the domestic or international press. If the domestic press, is the interviewer closely affiliated with the particular leader, more affiliated with that leader's opposition, or neutral in orientation? If the international press, to whom is the interview likely to be reported – people in an adversary's country, an ally's country, a country whose government the leader would like to influence, or a fairly neutral source? Of interest is whether the leader's trait scores show a pattern of change across these various types of audiences.

It is also important to consider if there have been any events (such as negotiations, crises, scandals, or international agreements) that have occurred during the tenure of the leader under examination. By noting when these events happened and choosing interview responses that span these points in time, it is possible to explore whether or not the leader's scores are affected by specific types of situations. In democratic societies, such an analysis might be conducted for periods before and after elections. For leaders with a long tenure in office, one might consider if there are any changes that have occurred across time with increased experience in the position or if the leader remains very much the same as when he or she began. Bill Clinton is an example (Hermann 2005; also Preston 2001; Preston and Hermann 2003).

If a researcher wants to assess mathematically whether changes in scores across time, topic, or audience are statistically significant, an analysis of variance will provide such data. Most statistical packages for personal computers have a one-way analysis of variance procedure, which can easily be applied to exploring this question. If the one-way analyses of variance (F-tests) are significant (have a probability value of 0.05 or less), then the leader's scores differ on that trait for that context factor (time, topic, audience); the leader is being adaptive in that type of situation.

Step 7: determining reliability of results

How easy is it to replicate the results of your content analysis? Can someone else using your coding rules and procedures end up with similar results to yours? There are two types of reliability that are often

calculated, and in examining the leadership styles of political leaders, I have found it important to ascertain both. The first assesses how easy it is for those unfamiliar with the coding rules and procedures to learn and apply them to leaders' interview responses with the same skill as that of the author. The second examines the stability in leaders' scores. This reliability is another way of ascertaining how sensitive a leader is to the political context.

Across a number of studies (Hermann 1980a,b, 1984; Hermann and Hermann 1989), the intercoder agreement for the traits used as examples here – belief that one can control events and need for power – have ranged from 0.88 to 1.00 between a set of graduate student coders and me. Where there were disagreements, the discussions that followed with the coders led to refinements in the rules and procedures. Generally, a coder is not allowed to content analyze a leader's interview responses included in the larger norming data set until he or she achieves intercoder reliabilities with me that are 0.90 or higher. These figures suggest that the coder can do almost as well as the author of the coding system in applying the rules and procedures. With the automated coding system Profiler + now available, similar types of reliability coefficients are being calculated to determine how accurate the software is in reflecting the original intent of the author of the coding system.

By correlating a leader's scores on odd- versus even-numbered interview responses, the researcher can ascertain how stable the traits under study are across time and issues. This reliability indicates how driven the individual leader is by his or her leadership style as opposed to responding to the situation. In effect, this index provides another way of determining how open or closed the leader is to contextual information (Hermann 1980a,b, 1984).

Step 8: ascertaining validity

How well does your analysis do in helping you learn about what you are interested in studying? In other words, how valid is your content analysis? Although there are a number of ways of ascertaining validity, four types are of interest here: content, predictive, concurrent, and construct validity. Another name for content validity is face validity and it 'is usually established through the informed judgment of the investigator. Are the results plausible? Are they consistent with other information about the phenomenon being studied?' (Holsti 1969: 143). Predictive validity deals with the ability to use what has been learned from the content analytic technique to forecast or understand future

events. Concurrent validity focuses on using the results. For instance, does what we learn via the content analysis about leadership style, say distinguishing between different types of leaders who exhibit particular styles, help us to differentiate among leaders whom historians and others have indicated challenged or respected constraints? Construct validity 'is concerned not only with validating the measure, but also the theory underlying the measure' (Holsti 1969: 148).

There appears to be a certain face validity to the at-a-distance assessment technique described above. It allows us to learn about what political leaders are like and to use somewhat similar procedures to those that are often employed to get at leadership style in the corporate world as well as in most counseling centers (see Nahavandi 2003). Moreover, it is possible to use what leaders say to infer certain characteristics that journalists, historians, biographers, psychologists, and political scientists have described as important to understanding political leadership. And political leaders appear to differ on these characteristics so that it is feasible to use the measures to compare and contrast the effects of leadership style not only on that person's behavior but on that of his or her government or other political unit.

A number of studies have been conducted to determine the predictive validity of the content analysis procedure described here. Generally the focus of this research has been to relate what we know about leadership style to what governments do. For example, one such study (Hermann and Hermann 1989) explored the behavior that leaders with different leadership styles urged on their governments. Among the findings were that those who challenged constraints were more likely to engage in conflictual and confrontational behavior, to take the initiative, and to commit the resources of their governments than those who respected constraints – results that were predicted from the literature exploring how leaders deal with constraints. Moreover, by comparing and contrasting the leadership styles of the various persons in positions of authority in Serbia during the Kosovo conflict, the analyst could understand why the crisis was being handled the way it was (Hermann 1999).

I have received numerous suggestions regarding how to assess the concurrent and construct validity of this methodology, ranging from running experiments with college students to participant observation in city councils. But it has always seemed important to find some means of comparing the results from the content analysis technique discussed in this chapter with the experiences of those who have interacted with heads of government and other prominent national leaders.

Thus, when given the opportunity to do such a project, I jumped at the chance to participate. In a series of studies, I developed profiles on 21 leaders from a variety of regions in the world following the rules and procedures described here and, based on these profiles, indicated on a set of rating scales the nature of the leadership behaviors the profile suggested that a particular political leader should exhibit (Hermann 1985, 1986, 1988). These ratings were compared with similar ratings made by people who had had the opportunity to observe or interact with these particular leaders.

At issue was how my results on leadership style could differentiate leaders recognized as different by journalists and former government personnel who had interacted with these leaders – concurrent validity. But it also tested the underlying logic and theory of the proposition that one could assess leadership style at-a-distance – construct validity. The correlations between the two sets of ratings averaged 0.84 across the set of 21 leaders, suggesting that the profiles derived from this at-a-distance measure furnished me with similar types of information on which to judge behavior as did the other raters' experiences with the actual figures.

Conclusion

The purpose of this chapter was to describe the steps and decisions involved in doing content analysis. I illustrated this method by using it to understand what political leaders are like. Content analysis provides us with a tool to gain some information about their beliefs, motives, and relationships with equals, subordinates, and constituents. It lets us take advantage of the fact that communication is an important part of what political leaders do. Indeed, archives are full of the speeches, press conferences, and writings of political leaders; the media seek to capture their interactions with political leaders on tape; and political leaders themselves often seek to preserve their legacies by building libraries around these materials or by donating them to universities and museums. By using materials from these various places, the technique described here is an unobtrusive measure of leadership style. And even though he or she may be shaping a communication for a specific audience or setting, we are able to take such intentions into account by varying the kinds of material we study.

Not only does content analysis make it feasible to construct a general profile of a particular leader or set of leaders but it is also makes possible placing such profiles into perspective by examining a number

of contextual factors that indicate how stable the characteristics are with certain kinds of changes in the situation. We can ascertain what leaders are like in general and, then, what kinds of information they are likely to be responsive to in the political environment. Thus, the general profile indicates where a specific leader fits among his or her peers and in a broader discussion of leadership style; the contextualized profile suggests how that leader has individualized his or her responses to manifest more unique characteristics. With knowledge about both the general and the more individualized profiles, the researcher and analyst gain a more complete portrait of a leader. The person becomes not only representative of a particular type of leader, but we know when and to what degree he or she has modulated their behavior to take the context into account.

It is important to remember that in doing content analysis, one's research question should drive the choices that are made regarding how qualitative one is going to be in designing the rules and procedures to be followed during the analysis, whether or not the material will be viewed as representative or instrumental, the nature of the assumptions to be made, and the choice of units of analysis. But if exploring what leaders are like and their effect on what governments do is a part of that research question, there are a growing number of content analysis techniques available to the researcher. In addition to those discussed here, the reader should consult Jerrold Post's (2005) compendium of these techniques and consider the types of analysis that are available through the computer software program Profiler + (see socialscienceautomation.com).

11

Pragmatic Analysis

Gavan Duffy

Pragmatic analysis refers to a set of linguistic and logical tools with which analysts develop systematic accounts of discursive political interactions. They endeavor to identify the full range of inferences that a reader or a hearer would make when encountering the locutions of an author or a speaker, considered in context. Consequently, pragmatic analysis is suited to the practice of inquiry that Hall (1999: 210–16) terms ‘specific history,’ in which analysts reconstruct, through emplotment, historical episodes that were meaningful to historical actors before they became meaningful as objects of analysis. Analysts endeavor to recover this meaning in order to understand agents’ actions and thereby to understand why events turned out the way they did rather than some other way.

The method is systematic, in the sense that any researcher may replicate the analysis of another. The method forces analysts to specify each inference explicitly. Critics who wish to dispute the substantive conclusions can point to the specific inferential steps that ostensibly misled them. This feature of pragmatic analysis focuses scholarly dispute on the source of intellectual disagreement, thereby promoting knowledge cumulation. Because it systematically examines meanings in the context of interaction, pragmatic analysis can be useful in empirically testing constructivist formulations, particularly those that theorize the role of agents in the creation of meanings, practices, structures, and institutions through their speech acts and communicative interactions.

Pragmatic analysis is a relatively new technique for analyzing political discourse. It arose from the Relatus Natural Language Understanding System, implemented by John Mallery and myself. Relatus is a collection of software tools that converts English text into a network representation in which nodes that represent concepts (or individuals) are

related to other such nodes. Relatus arranges these networks in such a way that they support various inferential procedures. After parsing and representing a narrative, the model could respond to our queries in English about the narrative's context. Hurwitz (1989) applied Relatus in analysis of prisoners' dilemma game play. It was also used to induce general if-then production rules from Frank Sherman's interstate conflict event dataset, SherFACS (Unsel and Mallery 1991; Sherman *et al.* 1992; Mallery and Sherman 1993; Sherman 1994).

Despite these successes, we recognized a deficiency in Relatus that would sharply limit its social and political applicability: it could not handle figurative language. As political language is replete with metaphors and other tropes that convey meaning implicitly, beneath the surface of a text, we regarded this limitation as fundamental. We first sought guidance in the continental literature on hermeneutics, as it had long been associated with interpretive approaches to social science. We soon discovered that this literature, however suggestive, did not provide the operational guidance we needed. It speaks in vague metaphors about 'effective histories' and offers figurative language about 'the fusion of interpretive horizons' by which humans convey meanings to one another (Hurwitz *et al.* 1987). We required more specific, procedural guidance, which we found in the literature on linguistic pragmatics.

Because it concerns context-specific inferences, pragmatics does not lend itself immediately to computer automation. Conversely, because computer models lack the rich, contextual experience required for sensitive, nuanced interpretations, they do not lend themselves immediately to pragmatic analysis. Consequently, pragmatic analyses have been non-computational, pen-and-paper affairs, at least initially, and only relatively short texts have been analyzed (Duffy *et al.* 1998; Frederking 2000; Duffy and Goh 2007). Once we gain sufficient experience in making the implicit explicit, we hope to 'recomputationalize' the technique. This should allow us to analyze larger textual corpora. In the interim, however, non-computational applications are feasible, so long as texts under analysis remain manageably sized.

Below, I review step by step the procedures of pragmatic analysis. I first provide an overview of the operations of pragmatic analysis, describing each step in a typical analysis. Next, I identify the grounding of these components in linguistic and logical theory. Along the way, I will illustrate by reference to an existing analysis of a relatively brief interaction between Henry Kissinger, Mao Zedong, and Zhou En Lai (Duffy and Goh 2007). The chapter concludes with discussions of the present limits of

pragmatic analysis and steps, computational and otherwise, that could be taken to extend them.

Operational overview

The flow chart in Figure 11.1 depicts the stream of operations in pragmatic analysis. I refer in the text to nodes in the figure by indicating

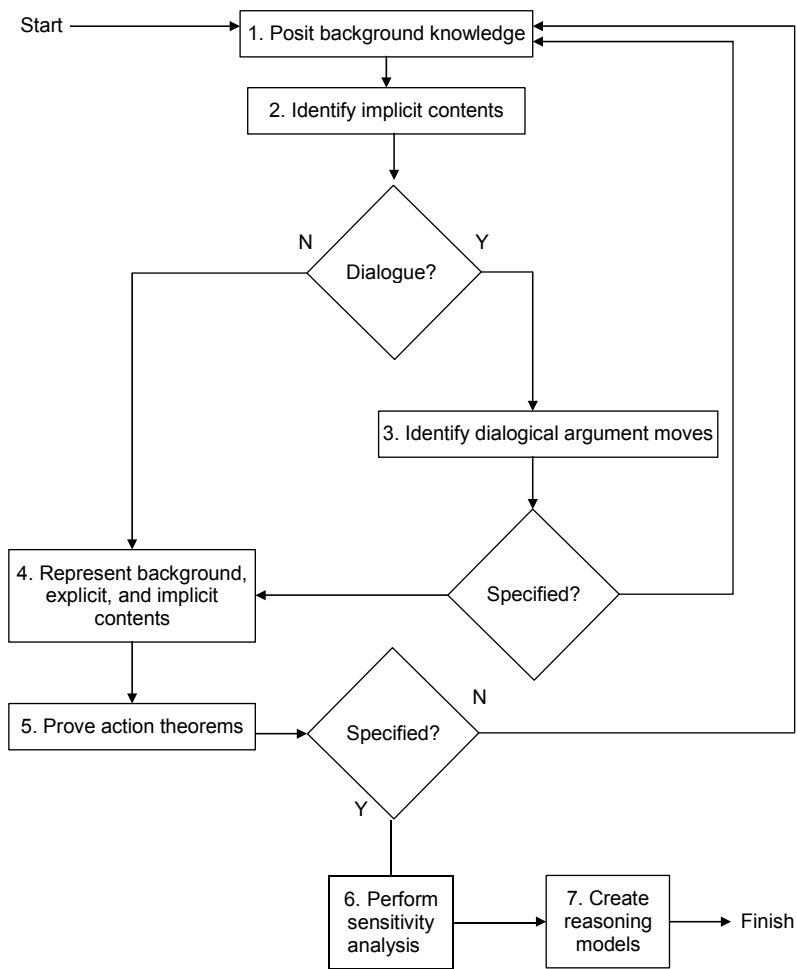


Figure 11.1 Flow chart of pragmatic analysis procedures.

the number of each within brackets. Space limitations prevent a full presentation here of the conversations between Kissinger, Mao, and Zhou, so examples will concentrate on just one utterance of Mao. Analyses of all the utterances in the conversation from which this line is drawn are available on-line at <http://web.syr.edu/~gavan/pragmatic/testing sincerity-pragmatic-analysis.doc>.

Pragmatic analysts first posit a set of background assumptions needed to interpret the text [1]. In hermeneutic terms, analysts specify a 'preunderstanding' necessary for interpreting the conversation. To avoid circularities in the analysis, these should be non-controversial. For instance, in modeling a conversation between Kissinger and Mao, it would be non-controversial to posit that Kissinger represents the United States and that Mao represents China, that China and the United States are states, and that states are sovereign. If the analytic interest were in modeling Kissinger's interpretation, then analysts might include propositions that are controversial, but non-controversial on (the analyst's understanding of) the mind-set of Kissinger. For example, analysts might non-controversially categorize Mao as a communist, but they might also add controversial propositions concerning the implications of communism, so long as evidence can be marshaled supporting the notion that Kissinger believed them, as we might glean from his writings.

Analysts next make their initial specification of the text's implicit contents [2]. They do so by explicitly drawing the inferences that make the text interpretable, given the background context. I should emphasize that these inferences, which I describe more fully in the section 'Linguistic Pragmatics' below, are entirely mundane. We are unused to stating them explicitly, but we all make them every time we read a text or hear someone speak. In our INF study (Duffy *et al.* 1998), for example, we made two inferences from the Soviet Union's offer during the INF negotiations to limit its missile deployment on condition that the United States cancel a planned deployment of its own: (a) that the Soviet Union wanted the United States to cancel its deployment and (b) that the Soviet Union believed that the United States wanted the Soviets to limit its deployment. In the preceding two sentences, I conveyed to you implicitly that these inferences would not prove difficult for you to draw were you to employ pragmatic analysis in your research.

If analyzing a debate or negotiation in which multiple parties are in conversation, one next conducts a dialogical argument analysis [3]. This step, described more fully in the 'Argument Analysis' section below, describes alternative moves and countermoves in the argument. Each such move or countermove is also a proposition stated explicitly in the

text or inferred analytically in Step [2]. Analysts record as additional implicit contents any propositions that a party to the dialogue commits herself tacitly or is driven by her interlocutor to concede tacitly. Often, when preparing an argument analysis, the analyst discovers that additional background knowledge or additional inferences are needed in order to render the argument interpretable. In this case, the analysis returns back to Step [1].

Once the argument analysis is complete – or if there is no dialogical argument and thus no Step [3] – the analyst represents the background knowledge, the explicit contents, and the implicit contents in a semantic network representation [4]. This step, along with those that follow, we envision for computational analysis. With small corpora, we can perform these operations by hand.

Once these contents are represented, we posit an ‘action theorem’ – a proposition that expresses the outcome of the dialogue or the conclusion with which the author wants the reader to agree. In the INF negotiation, for instance, the action theorem was simply that the Soviets wanted to reach an agreement. Analysts then prove that theorem follows logically from the body of beliefs represented in the semantic network [5]. If unable to prove the action theorem, the analyst concludes that the model is underspecified and searches for the missing background knowledge [1] or implicit contents [2] that render the action theorem consistent with the represented beliefs.

By representing in semantic networks the background context, the explicit contents of the text, and implicit contents inferred pragmatically and in the argument analysis, we have created a model that makes explicit all the contents that had originally been implicit and left for others in the conversation to infer from the context of utterance. By making all this explicit, it becomes possible to use formal logic to derive the action theorems. Without the pragmatic analysis, I should emphasize, the logical inferencing machinery would be blind to any context-specific inferences, of which the participants were well aware, owing to the context-sensitivity of their inferences.

Once the action theorems are proven, a sensitivity analysis is performed [6]. Here, propositions are retracted sequentially from the semantic network. If the action theorem remains consistent with the body of beliefs so represented, that proposition is deemed inessential to the practical reasoning leading to the action theorem. Once the inessential propositions are eliminated the remaining propositions are used to construct a syllogism [7] that models the analyst’s understanding of the reader’s interpretation of the text.

Note that we do not claim to have modeled the actor's interpretation, but only our interpretation of the actor's interpretation. Others may disagree with us, claiming that one or more of our premises or inferences is incorrect. However, because we have by then explicitly identified every single analytical inference we have made – because our analysis is replicable – anyone who wishes to challenge our conclusions can point to the specific premises or inferences that putatively led us astray. In other words, pragmatic analysis is a qualitative tool that is analytically rigorous because it produces replicable models. Statistical models are rigorous in the same way. They force the analyst to specify every inference, so that criticism is focused on the source of analytical disagreement.

I will now redeem my promises to supply additional details on the identification of implicit contents [2] and the identification of argumentative moves [3].

Linguistic pragmatics

The Anglo-American tradition of linguistic pragmatics offers specific guidance concerning the inferences that readers make (and authors anticipate) as they confront a text (Gazdar 1979; Levinson 1983, 2000; Blakemore 1992; Mey 2001). Linguists generally conceive pragmatics as the level at which contextually conditional inferences are made. In other words, pragmatics concerns those aspects of meaning that context-free, truth-conditional semantics (formal logic) cannot capture. In one formulation, *pragmatics* = *meaning* – *truth conditions* (Gazdar 1979: 2). I offer here a brief account of pragmatic inference and its application.

The pragmatic facets of meaning consist of speakers' (or authors') intentions in context. Consider the utterance of a speaker (S): 'Do you know what time it is?' Two possible hearings are *reflected in*: (H1) 'Yes, I do,' and (H2) 'Yes, it is 1:30.' H1 only responds to the *narrow* meaning of the query, while H2 fully understands that S is asking for a report of the time of the day. This little example highlights the difference between sentence meaning and speaker meaning. Grice (1957) explained the distinction as a consequence of the intentionality of human communication.

Communication is intentional and is achieved by the hearer's reflexive recognition of the speaker's intention. It consists of a speaker causing a hearer to think or do something by motivating the hearer's recognition of the speaker's effort to cause that thought or action. For example, when I state that 'communication is intentional,' I seek to motivate you

to recognize that I make the effort to state it because I believe it and I want you to believe it too. When earlier I promised to link pragmatic analysis to its linguistic foundations – which I am now endeavoring to do – I wanted you to recognize that I felt obligated to do so as a consequence of my promise.

Grice characterized speaker meaning as:

S meant *a* by uttering *U* if and only if:

- i. S intended *U* to cause some effect *a* in hearer *H*; and
- ii. S intended (i) to be achieved simply by *H* recognizing the intention (i).

As interaction unfolds, cooperative *speakers and hearers* convey to one another inferences about intentions that they draw from one another. By conveying and inferring intentions, these discourse partners progressively generate a body of ‘mutual contextual beliefs’ (Bach and Harnish 1979), a discourse context that conditions subsequent intentional conveyances and inferences. In this essay, I draw some examples from our pragmatic analysis of the February 1973 diplomatic interaction between Kissinger and Mao (with one contribution from Zhou) that appears in Table 11.1. Our analysis (Duffy and Goh 2007) teases out the intentions these discourse partners convey to and infer from one another in this conversation.

Grice’s theory of meaning closely resembles Austin’s (1962) and Searle’s (1969) theory of speech acts, which holds that speakers intend their locutions to have illocutionary force on hearers. That is, in uttering an expression, the speaker intends to have some effect on the hearer. Assuming *uptake* (the hearer’s successful recognition of the speaker’s communicative intent), the utterance will have some (perlocutionary) effect on the hearer – usually but not always the effect the speaker intended. Austin and Searle each hold that, with a successful speech act, a hearer recognizes certain ‘felicity conditions’ of the speaker’s utterances. These refer to the requisite thoughts, feelings, and intentions implicit in the speech act. A promise, for instance, carries with it an implicit claim of the speaker’s sincerity. Promises are felicitous if the hearer believes the speaker’s implicit sincerity claim.

Categorizations of performative verbs are useful because they group together verbs that evoke similar illocutionary and perlocutionary forces. Many have proposed categorizations. I find the Bach and Harnish (1979) taxonomy most appropriate, as it grounds itself in a plausible

Table 11.1 Interaction between Henry Kissinger (HAK), Mao Zedong (MAO), and Zhou En Lai (ZHOU) (February 1973, Beijing)

HAK:	We both face the same danger. We may have to use different methods sometimes, but for the same objectives.
MAO:	So long as the objective is the same, we would not harm you and you would not harm us. And we can work together to commonly deal with a bastard. In the West you always historically had a policy, for example, in both world wars you always began by pushing Germany to fight against Russia.
HAK:	But it is not our policy to push Russia to fight against China, because the danger to us of a war in China is as great as a war in Europe.
MAO:	<i>What I wanted to say is whether or not you are now pushing West Germany to make peace with Russia and then push Russia eastward.</i> I suspect the whole of the West has such an idea, that is to push Russia eastward, mainly against us and also Japan. Also probably towards you, in the Pacific Ocean and the Indian Ocean.
HAK:	We did not favor this policy...
MAO:	The whole of Europe is thinking only of peace.
ZHOU:	The illusion of peace created by their leaders.
HAK:	Yes, but we will do our best to strengthen European defenses and keep our armies in Europe.
MAO:	That will be very good.
HAK:	We have no plan for any large reduction of our forces in Europe for the next 4 years.
MAO:	We should draw a horizontal line – The US-Japan-Pakistan-Iran-Turkey-and Europe.
HAK:	We have a very similar conception. There is a strong community of interest that is operating between us.
MAO:	What do you mean by community of interest? On Taiwan?
HAK:	In relation to other countries that may have intentions.
MAO:	You mean the Soviet Union?
HAK:	Yes, I mean the Soviet Union.

theoretical account of the inferential responses of hearers to speakers' utterances. Moreover, Bach and Harnish posit their taxonomy in the context of an account of speech acts that merges Grice's theory of meaning and his related theory of conversational implicature. Their most general categories are the following:

Constatives – The expression of a belief (such as assertions, denials, descriptions, attributions, suppositions).

Directives – The expression of an attitude regarding some prospective action of the hearer (such as requests, requirements, prohibitions, permissions).

Commissives – The expression of an obligation or a proposal to obligate oneself under certain conditions (such as promises, guarantees, invitations, offers).

Acknowledgements – The expression of (perfunctory or genuine) feelings toward the hearer (such as apologies, thanks, greetings, congratulations).

Bach and Harnish (1979: 53–5) include no account of felicity conditions, reflecting the then-emerging scholarly consensus that these can be subsumed wholly under Grice's (1957) theory of conversational implicature (Levinson 1983: 241). In uttering a constative, for instance, speakers express their own belief as well as their desire that the hearer form (or continue to hold) that belief.

Grice contends that speakers employ rational principles for the effective use of language in order to further their cooperative ends. This rationality allows hearers to infer contents implicit in utterances. Grice terms these inferences as 'conversational implicatures.' Hearers implicate implicit contents by relying on several maxims of conversation that jointly constitute the cooperative principle: *make your contribution such as is required, at the stage at which it occurs, by the accepted purpose of the talk exchange in which you are engaged* (Grice 1957: 47).

1. *The maxim of quality*

- i. Do not say what you believe to be false.
- ii. Do not say that for which you lack adequate evidence.

2. *The maxim of quantity*

- i. Make your contribution as informative as is required for the current purposes of the exchange.
- ii. Do not make your contribution more informative than is required.

3. *The maxim of relevance* – Make your contributions relevant.

4. *The maxim of manner*

- i. Avoid obscurity.
- ii. Avoid ambiguity.
- iii. Be brief.
- iv. Be orderly.

Conversational implicatures preserve the principle of cooperation because hearers assume that the maxims are being observed.

Speakers do sometimes fail to observe a maxim. But, unless acting irrationally, they do so with illocutionary intent. A flouted maxim signals to a reader/hearer that the speaker/author does not intend the literal meaning of the utterance. Comparing the utterance to the store of mutual contextual beliefs, the hearer tests whether the speaker intends sarcastic, ironic, metaphorical, or exaggerated meaning (Bach and Harnish 1979: 68–9). For instance, a prolix apology issued to someone who incorrectly believes herself the victim of a minor social slight, for instance, violates submaxim (iii) of the maxim of manner, to be brief. This violation allows the hearer to infer that the speaker is not apologizing at all. The discourse context usually contains sufficient clues as to the speaker's intent, because the speaker (being cooperative) expects to be understood by the hearer. In the unusual circumstance that the hearer does not hit upon the speaker's illocutionary intent, the speaker will notice (by the hearer's inappropriate response) the failure of uptake and take steps (perhaps by being more explicit) to repair the hearer's understanding of the speaker's intent.

In strategic political dialogues, participants sometimes say what they believe to be false. That is, they sometimes flout submaxim (i) of the maxim of quality for strategic purposes and not to signal non-literal communicative intent. Thus, in analyzing the talks between the United States and China, we view with some skepticism any implicatures that arise from this maxim, especially in the presence of other evidence that the participants are not communicating cooperatively. For instance, we consider Mao's statement, 'What I wanted to say is whether or not you are now pushing West Germany to make peace with Russia and then push Russia eastward,' to be an indirect speech act. We characterize it as formal locution (A) in Table 11.2. ('Formal locution' is our term for a restatement of the surface – or actually spoken – utterance to resolve pronoun antecedents and other indirect locutions.) Although it appears to be an expression of belief (constative) on the surface, Mao would have violated the maxim of relevance were he not indirectly asking Kissinger whether US policy toward West Germany was directed toward encouraging the Soviet Union's attention eastward. As an indirect question, Mao's proposition suggests reflexive intentions (C) and (D). Mao wants Kissinger to answer the question, stated explicitly in (B) and that he wants Kissinger to do so.

Implicature (G) helps to bring Mao into conformity with the maxim of relevance by noting that he asks this question because Kissinger's response to Mao's previous statement, 'But it is not our policy to push Russia to fight against China, because the danger to us of a war in China

Table 11.2 Analysis of italicized proposition in Table 11.1

What I wanted to say is whether or not you are now pushing West Germany to make peace with Russia and then push Russia eastward.

Formal locution

- A. Is the US now pushing West Germany to make peace with the SU in order to push the SU eastward?

Indirect Speech Act: (Question, Directive)

Explicit Performative

- B. Mao hereby asks (A).

Reflexive Intentions

- C. Mao wants HAK to tell Mao whether the US is now pushing West Germany to make peace with the SU in order to push the SU eastward.
 D. Mao wants HAK to tell Mao whether the US is now pushing West Germany to make peace with the SU in order to push the SU eastward at least partly because of Mao's desire (C).

Presuppositions

- E. The US can push West Germany to make peace with the SU and then push the SU eastward.
 F. If West Germany makes peace with the SU, the SU is free to turn its attention to its eastern front.

Implicatures

- G. In (6), HAK failed to recognize Mao's intention in (5). (*5 and 6 refer to propositions earlier in the conversation*)
 H. Mao believes that by pushing West Germany to make peace with the SU, the US indirectly pushes the SU eastward.
-

is as great as a war in Europe,' indicated a failure of uptake on Kissinger's part. Mao's locution, 'what I wanted to say...', indicates that Mao found inappropriate Kissinger's response to Mao's reference to a putative Western policy in both world wars of pushing Germany eastward.

Implicature (H) concerns Mao's 'and then' locution. Ordinarily, this locution signifies a temporal interpretation. That is, a locution of the form 'x and then y' indicates that x temporally precedes y. In this case, however, in order to conform Mao's utterance to the maxim of relevance, the conjoined clauses represent a process – the United States pushes West Germany to make peace with the Soviets in order to push the Soviets eastward.

Our analysis of Mao's utterance also includes two presuppositions: (E) that the United States can perform the action about which Mao asks, and (F) that the Soviet Union can turn its attention to its eastern front. A presupposition of a statement is a proposition that must be true in order for that statement to have a truth-value. A speaker/author

implicitly commits to the truth of the presuppositions simply by uttering the statement. For instance, consider Kissinger's assertion 'We did not favor this policy,' referring to Mao's contention that the United States sought to create trouble for China on its western front. Kissinger here presupposes that such a policy did exist (or was floated).

Presuppositions differ from entailments because they are constant under negation. If one negates Kissinger's statement so that it reads, 'We did favor this policy,' the inference remains. If this were an entailment, negating the statement would negate the inferred proposition. But it does not. Presuppositions are also non-cancelable without anomaly. Kissinger could not have said (without anomaly) 'We did not favor this policy, which did not exist.' If the prospective policy did not exist, then the question of whether to favor it would not have arisen. Implicatures, on the other hand, may be cancelled without anomaly (for instance, 'I promise to be truthful, but I'm insincere').

Many varieties of presupposition have been identified. Consider the factive verbs, such as 'regret': 'The Soviets regret shooting down KAL 007' presupposes that 'The Soviets shot down KAL 007.' Changes of state sometimes signal presuppositions. 'Germany ceased its bombardment' presupposes that it had been bombarding something. Levinson (1983: 179–85) reviews a number of 'presupposition triggers,' or linguistic forms that signal the presence of a presupposition. (Efforts to provide a semantic theory of presupposition have ultimately failed. The 'projection problem' presents an insurmountable obstacle to anyone interested in mechanizing presuppositional inferences and motivates a general linguistic preference for pragmatic over semantic accounts of presupposition. See Levinson 1983: 185–98.)

Although we do not use them in our example, scalar implicatures appear often enough to be mentioned here as a source of implicit content. These consist of inferences that hearers/readers make to bring speakers/authors into conformity with the maxim of quantity. There are a large variety of such scales, including <all, most, many, some, few>, <hot, warm>, <succeed in V-ing, try to V, want to V>, <n, ... , 5, 4, 3, 2, 1>, <love, like>, <always, often, sometimes>, and <necessarily *p*, *p*, possibly *p*>. For example, consider the scale <must, should, may>. As one replaces *x* with one of these modal operators in the proposition 'Iraq *x* withdraw from Kuwait,' then propositions with operators to the right of *x* are scalar implicatures of the proposition and thus count as implicit contents. So 'Iraq must withdraw from Kuwait' implicates 'Iraq should withdraw from Kuwait' and 'Iraq may withdraw from Kuwait.' Likewise, 'Iraq should withdraw from Kuwait' implicates 'Iraq

may withdraw from Kuwait.' Note that we do not implicate in the other direction. 'Iraq may withdraw from Kuwait' implicates neither that it should nor that it must. It is important to remember that scalar implicatures, like all implicatures, are cancelable in context.

Deictic inferences concern the relative distances between objects in the text, as indicated by pronominal and other indexical expressions (such as *here, there, now, then*). Wilson (1990) analyzed deictic expressions in (the then) British Prime Minister Margaret Thatcher's speech. In addition, political actors can convey content by means of prosodic stress and intonation. If audio or video recordings are available, an analyst would be well advised to pay attention to verbal signals and gestures as additional sources of implicit content.

Argument analysis

We borrow and modify slightly Rescher's (1977) models of alternating and context-elaborating flow of argumentative dialogue. Patterned after the medieval procedures of thesis defense, Rescher has the proponent(s) of a thesis assert categorical or provisoed claims to the opponent(s), who skeptically issue(s) cautious or provisoed denials of these claims. He expresses provisos in the form P/Q , read as ' P ordinarily, if Q ,' or 'if Q then, *ceteris paribus*, P .' In his notation, Rescher prefixes the categorical claims of the proponent with an exclamation point (!). He prefaces the cautious claims of the opponents with a dagger (\dagger). Table 11.3 presents a hypothetical illustration.

Table 11.3 Hypothetical illustration of Rescher's Formalism

Proposition	Proponent	Opponent
1. P is the case.	!P	
2. Please show P to be the case.		$\dagger \sim P$
3. P is ordinarily the case when Q and I assert Q.	$P/Q \ \& \ !Q$	
4. P ordinarily isn't the case when Q and R and, for all you've shown, Q and R could be the case.		$\sim P/(Q \ \& \ R) \ \& \ \dagger(Q \ \& \ R)$
5. R is ordinarily not the case when T, and I assert T.	$\sim R/T \ \& \ !T$	
6. P is ordinarily not the case when U, however, and for all you've shown, U could be the case.		$\sim P/U \ \& \ \dagger U$

In the process of the argument, proponents accumulate commitments to particular propositions and opponents concede particular propositions. By Move 6, for instance, our hypothetical proponent is committed to propositions P , Q , $\sim R$, and T . The opponent, meanwhile, tacitly concedes Q by responding in Move 4 with $P/(Q \& R) \& \uparrow(Q \& R)$ instead of $\uparrow \sim Q$. Of course, the opponent may subsequently retract this concession by denying (with or without proviso) Q in later moves. As these commitments and concessions accrue, they become implicit beliefs of the proponent or opponent.

In applying Rescher's formalism to Thucydides' Melian dialogue, Alker (1988) showed it to be capable of modeling political argument. Nevertheless, some modifications of Rescher's formalism are needed. In particular, our use requires that we enable the discourse partners to shift to one another the burden of proof. In the thesis defense scenario, the proponent always bears the argumentative burden. However, because burden-shifting is a pervasive argumentative strategy in modern discourse (Gaskins 1992), we expect the burden to shift frequently across participants, especially in strategic settings. For this reason, we modify Rescher's formalism to allow either side to advance new propositions and to deny propositions advanced by their adversaries.

The argument analysis of the Kissinger – Mao conversation appears in Table 11.4. Associated with each proposition is a cognate from the speech act analysis of the dialogue (found on the webpage referenced above). In Proposition 3, Mao issues a provisoed denial ($\sim S/F \& !F$) of Kissinger's contention (S) that China and the United States share objectives on the grounds that the United States is encouraging the Soviets to attack China (F). This encouragement, Mao further contends in Proposition 4, means that China and the United States cannot maintain peaceful relations. Kissinger responds in Proposition 5 with a provisoed denial of F ($\sim F/D \& !D$). The United States would not encourage a Soviet attack on China because a Sino-Soviet war would endanger the United States.

Mao responds in Proposition 6 (the example sentence analyzed in Table 11.2) with a provisoed assertion ($F/W \& \uparrow W$) that the United States is encouraging the Soviets to attack China indirectly, by pushing West Germany to make peace with the Soviets. Kissinger first responds in Proposition 7 by categorically denying ($!\sim W$) that this is US policy. He immediately supports this denial in Proposition 8, by promising (E) that the United States will maintain its forces in Europe. We interpret this as a provisoed assertion ($\sim W/E \& !E$). Note that in Proposition 9, Mao accepts this promise ('That will be very good,' he says), which we interpret as his acceptance of the provisoed assertion. Kissinger reinforces $\sim W$ in

Table 11.4 Argument analysis of the Kissinger–Mao interaction

Kissinger		Mao	
Proposition	Cognate	Proposition	Cognate
1. $!S$	2g	2. $\sim P/\sim S \ \& \ \dagger \sim S$	3k
		3. $\sim S/F \ \& \ !F$	5g
		4. $\sim P/F \ \& \ !F$	3h, 5g
5. $\sim F/D \ \& \ !D$	6a	6. $F/W \ \& \ \dagger W$	7h
7. $!\sim W$	10a		
8. $\sim W/E \ \& \ !E$	14k	9. $\sim W/E \ \& \ !E$	15e
10. $\sim W/M \ \& \ !M$	16f		
11. $!S$	19a		

S: The US and China share objectives; *P*: China can maintain peaceful relations with the US; *F*: The US is encouraging the SU to attack China; *D*: A Sino-Soviet war would be dangerous to the US; *W*: The US is pushing West Germany to make peace with the SU and push the SU eastward; *E*: The US will strengthen its European defenses and keep its armies in Europe; *M*: The US does not plan any large reduction of its forces in Europe for the next 4 years.

Proposition 10 with another provisoed assertion, $(\sim W/M \ \& \ !M)$, which conveys that the United States plans no near-term force reductions in Europe. Then Kissinger restates in Proposition 11 the assertion with which he opened the interaction, that the United States and China share a common interest ($!S$). In the remainder of the interaction, unanalyzed here, Kissinger and Mao go on to clarify that this common interest extends only to relations with the Soviet Union, and not to the issue of Taiwan.

Due to the relative brevity of the Kissinger–Mao–Zhou interaction, we need not create computational models, derive action theorems, and perform sensitivity analyses in order to generate a practical reasoning model. We can do this by hand, as Mao’s argument is relatively straightforward and simple. The propositional logic model in Table 11.5 succinctly expresses our understanding of Mao’s reasoning. It supports the action theorem we posit, which denies the sincerity of Kissinger’s claim that the United States and China share objectives ($\sim S$). In Step 4, from his belief that the United States is pushing West Germany to make peace with the Soviets (W), Mao infers that the United States (indirectly at least) is encouraging a Soviet attack on China (A). From this, Mao infers in Step 5 that Kissinger’s overture is insincere.

Table 11.5 Model of Mao's reasoning

1.	W	premise
2.	$W \supset A$	premise
3.	$A \supset \sim S$	premise
4.	A	1,2 MP
5.	$\sim S$	3,4 MP

W : US is pushing W. Germany to make peace with the Soviet Union; A : The US is encouraging the Soviet Union to attack China; S : HAK, when claiming that the US and China share objectives, is sincere.

On this analysis, Kissinger can establish sincerity only by dissuading Mao of the validity of these inferences. He does this, first, by observing that a Sino-Soviet war would pose dangers for the United States. Shortly thereafter, he reinforces this by promising to maintain and even bolster European defenses. These steps effectively deprive Mao of the premise $W \supset A$, and thus Mao can no longer infer $\sim S$, that Kissinger's overture to the Chinese leaders is insincere. Duffy and Goh (2007) rely on this demonstration to show that these leaders in this real-world context of strategic interaction test the sincerity of other actors' utterances by assessing their consistency with those actors' interests – just as has been suggested for pragmatic analysis (Duffy *et al.* 1998) and for game analyses that integrate verbal interaction with game models.

We analyze the Mao-Zhou-Kissinger conversations in order to discover how Kissinger's plan to engineer a tripartite equilibrium between the United States, China, and the Soviet Union eventuated in what amounted to a secret alliance between the United States and China. Our analysis of this early conversation cannot address this question, but it does provide some insights into the specific history of the interaction. These raise questions that could inform our subsequent analyses of later conversations. Most importantly, did Chinese skepticism of Kissinger motivate Kissinger to offer China what amounted to a secret alliance – intelligence on Soviet troop movements, establishment of a hotline to inform China of US early warning information, and other forms of material and technical military assistance?

The scope of pragmatic analysis

Every analytic technique has a scope – a range of questions over which it can and cannot find application. Pragmatic analysis is no exception.

It is inappropriate for the practice of inquiry that Hall (1999) terms 'analytical generalization,' and should not be judged on those terms. Pragmatic analysts do not test hypotheses derived from general theoretical propositions by examining the associations of operational variables to see whether they associate as theory predicts. Neither do they conduct experiments or quasi-experiments in which they examine how different treatments (or lack of treatment) affect experimental subjects, be they individuals or polities. However widely read their views on qualitative method may be, King, Keohane, and Verba (1994) would inappropriately restrict the range of inferences appropriate for political analysis.

Although pragmatic analysis appears just as constructivism is rapidly advancing in international studies, it cannot readily receive application to all questions in the constructivist vein. As Crawford (2002) discovered, the method is simply impractical for addressing questions of world-historical scope. This limitation notwithstanding, pragmatic analysis nonetheless has utility for constructivist research projects. '[T]he minimum we should expect of any effort to test constructivist claims,' writes a prominent liberal institutionalist critic of constructivism, 'is not just the derivation of fine-grain empirical predictions ... but also the utilization of methods capable of distinguishing between spurious and valid attributions of ideational causality' (Moravcsik 1999: 675). Although I prefer to speak of 'intentionality' over the clumsy 'ideational causality,' pragmatic analysis – by reconstructing the intentions of political agents – is a method capable of making such distinctions.

In the Netherlands, van Eemeren and Grootendorst (2004; van Eemeren *et al.* 1993) have developed argument analysis techniques that share some features with the method described here. Their 'pragmadiialectical approach' draws upon speech act theory to analyze argument moves and countermoves. The philosophical and theoretical underpinnings of their approach broadly cohere with those of pragmatic analysis. There are several important differences, however. For instance, van Eemeren and Grootendorst do not analyze argument commitments and concessions as in Rescher's formalism. Neither do they conduct sensitivity analyses to produce the core of arguments expressed in logical notation. Owing to their practical interest in improving the reasonableness of argumentation across a wide range of venues, van Eemeren and Grootendorst posit a normative model of reasonable argumentation and show how particular arguments have deviated from that norm. Pragmatic analysis has thus far been deployed to study the arguments of political elites. Because political elites tend to be highly educated

(or at least surrounded by highly educated advisers), their arguments (with each other, at least) tend to conform to the norms of reasonable argumentation. These differences aside, anyone interested in conducting a pragmatic analysis would be well advised to examine the work of van Eemeren and his associates, as their numerous insights into argument modeling can usefully inform any such project. Close attention to this work is warranted should one wish to analyze the arguments of non-elites.

In the development of any new technique, it makes a great deal of sense to begin with relatively small, more tractable problems. Once they gain experience using the technique, developers can then scale it up to address larger problems. Pragmatic analysis can be scaled up, to some extent at least, by automating steps in the technique. The theorem-proving tasks – proving the action theorems and conducting sensitivity analyses – can readily be automated, as they are context-independent operations. Techniques for theorem proving are well known and a prototype has already been implemented in a Relatus-like knowledge network.

Because their operations are context dependent, it will not be feasible to automate fully earlier steps in the analysis. The argument analysis step, for instance, cannot in principle be automated, as decisions regarding the propositions to select for each argumentative move depend entirely upon interpretations of the context that only humans can perform. I suppose one day these operations might be automated, but mechanizing them presumes a solution to the Artificial Intelligence problem – the provision of an artificial intelligence indistinguishable from human intelligence. Until we get there, we should stick with human intelligence.

Prospects for automation are rosier where speech act analyses are concerned. Levinson's (1983) discussion of presupposition, for instance, reviews a dispute that raged between linguists who advocated a pragmatic account of presupposition and linguists who advocated a semantic account. The semantic account proposed a context-independent theory of presupposition that was modified each time an advocate of the pragmatic account found some new context-dependent anomaly the semantic theory could not handle. In the end, Levinson casts his lot with the pragmatic account, as context-dependent counter-examples kept cropping up, no matter how many complexifications semantic advocates added to their theory. However much it may have failed to offer a context-independent theory of presupposition, the semantic theory offers crucial guidance for the design of a program that would suggest potential presuppositions to a pragmatic analyst.

Likewise, straightforward programs could be written that would suggest speech act categories and reflexive intentions to an analyst, who could accept them or – if a speech act were indirect or if the main verb were used in an unusual sense – override them. The crucial point here is this: although the context-dependent operations of pragmatic analysis cannot in principle be fully automated, they could usefully be partially automated as a toolkit that analysts could use as they ‘code’ or interpret an interaction. Pragmatic analysis may never satisfy Crawford’s (2002) desire for a method capable of analyzing world-historical trends over several centuries. But it is not difficult to envision scaling the method up to support analyses of larger corpora than the relatively brief exchanges we have thus far analyzed.

Conclusion

I close with a plea. Whether or not pragmatic analysis is appropriate for analyzing the substantive problems that interest you, the pragmatic theory on which it relies should interest anyone pursuing familiarity with qualitative methods. Most of these, if not all of them, involve the examination of meanings that people produce in social contexts. Pragmatic theory is our best theory of meaning in social context. As such, pragmatic theory serves for qualitative methods the same foundational role that probability theory serves for quantitative methods. We can get along fairly well without awareness of this relationship, as pragmatic theory concerns inferences that we all make routinely as we exchange meaning-contents with others in our everyday lives. However, reflection on the nature of exchanges of meaning-contents can make us better qualitative analysts. Studying and understanding pragmatic theory – even if not explicitly practicing pragmatic analysis – yields that reflection.

12

Agent-based Modeling

Matthew J. Hoffmann

This book on qualitative methods fills a significant gap in the international relations methods literature by avoiding the attempt to create or reify orthodoxy around a particular set of methodological tools. The book also addresses a second problem with the methods literature: the lack of serious discussion of the multiple methods of qualitative analysis, including a comparison of their kind of scholarship or knowledge production. Of course, given the goals and title of this book, one has to ask, Who let in the formal modeler?

This chapter discusses a relatively peculiar method for inclusion in a book on qualitative methods: agent-based modeling (ABM). While a number of applications fall under the ABM umbrella, the unifying goal of this method is to simulate and understand processes through which macro- or social patterns emerge from the actions and interactions of agents (and their context). While well known in other social sciences (Parker *et al.* 2003; Macy and Willer 2002), ABM has only recently made inroads in international relations (Harrison 2006; Johnson 1999; Rousseau and van der Veen 2005; Majeski 2004, 2005; Simon and Starr 2000). Robert Axelrod (1984, 1997) introduced it with his prisoner dilemma tournaments; he has also simulated the growth of empires and tribute systems. Lars Erik Cederman (1997, 2001, 2003) has produced an impressive array of models that capture the emergence of nations and the distribution of wars over time. Ian Lustick (2000) and his collaborators (Lustick *et al.* 2004) have worked with cultural transmission and political identity. Recently I modeled the emergence and evolution of social norms in environmental politics (Hoffmann 2005a).

The inclusion of ABM in this book is potentially curious because most of those who use the tool would likely **not**, at first thought, consider it to be a qualitative method. Yet, its inclusion here is entirely appropriate

if only because it allows us to problematize the perceived distinction between formal and qualitative methods. Yes, ABM is a formal method. It involves writing computer programs (formal algorithms) to simulate social/political behavior. Yes, ABM could technically be considered a quantitative method in that numbers are used to represent political concepts. However, beyond these surface observations we can note characteristics of ABM that confound easy classification. While the method does begin with formal algorithms, it does not deductively prove theorems or reach closed form analytic solutions, as do the more familiar formal methods like game theory. In addition, while ABM is numerical in important senses, it does not operationalize political concepts into numerical variables in order to test statistical hypotheses like econometrics.

Instead, ABM is more of a social laboratory and, depending on the model written, shares more in common with rich process tracing case studies or metaphorical analysis than it does with game theory or econometrics. If we consider that qualitative analysis aims at understanding the qualities and meanings of political relationships (while quantitative analysis aims at hypothesizing about the implications of varying quantities of political variables), then ABM is indeed a qualitative method, despite its formal and numerical characteristics.

The rest of this chapter defends the claim that ABM belongs in this book and in the toolkit of international relations. My main purpose is to introduce the general idea of ABM, along with a set of guidelines to enable someone to answer the question: Should I learn to model? I begin by providing the basic philosophy and practical steps for constructing ABM simulations. With an example from my model of social norm dynamics, I then show the method used to approach social facts usually dealt with by constructivist case study analysis. I follow this with a discussion of what kind of knowledge is produced in ABM exercises and its epistemological and practical implications (both good and bad). While the chapter does make a case for ABM as an appropriate tool for studying world politics, I do not advocate modeling in the absence of additional empirical methods.

Methodology: what is simulation modeling?

In principle, the logic of ABM is very simple. The construction of a simulation model is encompassed in four steps (see Epstein and Axtell 1996; Conte *et al.* 1997; Johnson 1999; Page 1999). The essence of ABM lies in the creation of artificial agents. These can be envisioned

as individuals, organizations, or even states depending on the research question. 'Creating' an agent consists of endowing computer objects with: (a) characteristics, like age, identity, and behavioral rules, that can change from simulation to simulation; (b) the ability to perceive the environment, such as reception of information about their surroundings and/or fellow agents; and (c) a decision-making apparatus, providing the ability to take actions.

Second, the modeler defines the artificial environment within which the agents interact. This environment can be wholly social, consisting only of interaction rules that govern how agents relate to one another. The environment can also have spatial and physical (though still artificial) characteristics. Agents can be placed on grids, with each having a set location or the ability to move (Lustick 2000). Alternatively, they can be placed in a soup with no set location and random interactions with each other (Axelrod 1997; Hoffmann 2005a). Agents can be placed on a representation of an actual landscape, for instance, marrying geographic information systems (GIS) with ABM. The environment itself can be merely a place for agent interactions, or it can have characteristics of its own like resources and growth dynamics (Epstein and Axtell 1996; Hoffmann *et al.* 2002; Janssen 2002; Parker *et al.* 2003).

The third step is simulating histories in this artificial world. The modeler initializes the agents and environment with a set of conditions and then lets the agents (or agents and environment) interact for a specified amount of time. These simulations are usually calculated in rounds of actions taken by the agents or time-steps that may or may not be analogous to real-world time steps. As the simulation history progresses, data are gathered.

Finally, the modeler analyzes the history through visual (graphic) or statistical means. This analysis is done to ascertain what happened in the simulation, that is, how the initial micro-parameters correspond with the emergent macro-parameters, and/or how the simulation history matches empirical observations. In this stage there are significant concerns with validation, both internal and external. (For more on this much debated topic, see Oreskes *et al.* 1994; Casti 1997; Parker *et al.* 2003.)

The key to ABM is its flexibility. It is possible to endow the agents and environment with almost any kind of attributes, decision-making rules, and interactions imaginable. For instance, agents can be heterogeneous and adaptive, rather than homogeneous and rational. (ABM can accommodate game theoretic and rational choice analysis, making game theory a special case of ABM, where closed form solutions are possible.)

No longer does a commitment to non-rational behavioral models or process-oriented theorizing require eschewing formal analysis. ABM's adaptive agents are often designed with limited computational ability, and its artificial environments are worlds of less than complete information (in the rational choice sense of completeness), which many theorists claim is a more accurate representation of reality than calculating the optimal course of action based on (often) full information of all alternatives. These adaptive agents 'rely on heuristics or rules of thumb' that are learned over time, through experience (Kollman *et al.* 1997: 465). In addition, they 'inhabit a world that they must cognitively interpret – one that is complicated by the presence and actions of other agents and that is ever changing'; they 'generally do not optimize in the standard sense...because the very concept of an optimal course of action often cannot be defined' (Arthur *et al.* 1997: 5).

This type of analysis defies easy classification. The deductive or formal aspect is the *a priori* designation of what actors look like and how they behave. The modeler creates agents as well as a context of interaction. As the goal is to understand some macro- or social phenomenon, the initial designation of characteristics or rules is essentially a hypothesis or conjecture about how micro-characteristics are related to macro-phenomena through a specified process. I consider the initialization as akin to 'cutting-in' to a dynamic process at a specific point to begin analysis with process tracing case studies.

Yet ABM is not solely a deductive approach. Once the rules and characteristics are initialized, the modeler simulates histories of the population of agents and their interactions, varying parameters of interest. Analyzing the results entails an inductive assessment of the simulated histories. Axelrod (1997) calls it a third way of doing science that is both deductive and inductive. Epstein and Axtell view the simulations as '*laboratories*, where we attempt to "grow" certain social structures in the computer – or *in silico* – the aim being to discover fundamental local or micro mechanisms that are sufficient to *generate* the macroscopic social structures and collective behaviors of interest' (1996: 4).

ABM's flexibility opens up possibilities for formalizing approaches to social life that have heretofore distanced themselves from rational choice. Applied to theories that highlight the importance of ideas, language, and interpretation, ABM can bring together classic qualitative approaches, like those represented in this book, with formal theory. To demonstrate this potential, I briefly discuss my application of ABM to social norms, which draws on the constructivist IR literature (Hoffmann 2005a,b).

A modeling example – the norm life cycle

Finnemore and Sikkink's (1998) norm life cycle provides a set of conjectures about norms in a society, including how they constitute agents, and how they emerge and evolve. Their framework essentially comprises a verbal model of norm dynamics. I model this more formally by (Step 1) endowing artificial agents with behavioral characteristics and (Step 2) setting up interaction rules, along the lines suggested by ideas of mutual constitution. I then (Step 3) let the agents interact creating simulated histories of norm dynamics, under varying conditions, and (Step 4) analyze the results to see what emerges. As I will discuss in greater detail below, this is a heuristic model designed to apprehend some generic features of the emergence and evolution of norms.

Agents in some constructivist theorizing are conceived of as striving to behave appropriately. Appropriateness is defined by the agents' social context. This context, in turn, is constructed by the behaviors and interactions of the agents. Norms are ideas about appropriate behavior that are intersubjectively held within a population. According to the Finnemore and Sikkink (1998), this norm life cycle encompasses four stages. First, a norm entrepreneur makes a suggestion to a population about a new way to conceive of appropriate behavior. When a 'critical mass' of actors have accepted the new idea, we can say a norm has emerged. In the third stage, there is a cascade of acceptance as the norm spreads from the critical mass to the whole population. Finally, after use and reification, the norm is internalized and becomes taken for granted.

I constructed a model to formalize this norm life cycle. It simulates 10 agents, an arbitrary but manageable number. Each are driven by the desire to match their individual behavior to that of the group, that is, to act appropriately. Their behavior is very simple. In each round of the simulation, agents pick a number between 0 and 100, an arbitrary range that provides significant room for variation in behavior. Their goal is to match this pick with the group outcome, which I have defined to be the average (arithmetic mean) of the choices from the entire group.

Many constructivists envision actors to be rule-driven agents, so my agents make their predictions with a universe of seven (very) simple rules. The rules drive the agents to pick numbers from bounded sets of numbers. The actual choice of number is drawn from a uniform distribution of integers within the specified boundaries: Rule 1, 0–10; Rule 2, 15–25; Rule 3, 30–40; Rule 4, 45–55; Rule 5, 60–70; Rule 6, 75–85; Rule 7, 90–100. In essence, the rule provides a broad sense of what an

agent thinks is an appropriate behavior (or number, in the simulation), and the agent picks a specific number from within that broad sense.

Initially each agent is randomly assigned three of these prediction rules (without repeats) in order to insure diversity in the population of agents. These initial conditions can be considered in two ways. One could begin in a state-of-nature, with a few hardwired rules with which agents approach social interaction. I prefer to view the initial rule assignments as akin to 'cutting in' to a dynamic process at a specific point in time, when the rule distribution looks as it does. Because I am not attempting to simulate a specific political situation or norm, the initial distribution of agent attributes is less significant than it is for empirical analysis, when it is necessary to understand how actors came to be socialized as they are when you begin to study them (the importance of which is highlighted in both the Leander and Checkel contributions to this book).

In line with at least some constructivist thought, these are reflective agents – they do not blindly follow the social context but actively consider what behavior is appropriate. Each agent uses one of its three rules – the public rule – to make the prediction that is sent to the entire group, while privately making predictions with the other two rules. The public rule generates the prediction to which the whole group of agents has access. Each agent determines which rule is public by keeping track of scores for the rules in its repertoire. Starting with a baseline of 100, each score rises and falls depending on how close its predictions have been to the group outcome. The rule with the highest current score is the public rule.

In order to judge satisfaction with their rules, the agents evaluate the behavior produced by the public rule as well as the potential behavior of their other two private rules. Once the group outcome is known, agents compare their three predictions (one public, two private) with that outcome. They reward or penalize their rules depending on the closeness of the prediction. In this model, 'close enough' is governed by a parameter called precision, set (in most runs) at 5 per cent. This means that rules which predict the group outcome within ± 5 are rewarded ($+1$) and others are punished (-1). A private rule becomes public when its score exceeds that of the current public rule. To facilitate adaptation and change over time, each agent discards a poorly performing rule at set intervals (10–20 rounds) and is randomly assigned a new rule from the universe of rules. The new rule starts with a fresh score of 100. Such shuffling of rules could be conceived of as domestic change within a state or shift within an international organization.

The agents' social context is limited in this simple model; they only perceive the group outcome. This characteristic is designed to mimic the limited sociality of world politics. It should be noted that the model does not explicitly simulate how agents obtain an understanding of their social environment (such as through communication and other social activities). Instead, it focuses exclusively on what happens once agents have a picture of their social environment. This choice allows me to concentrate on how an entrepreneur can catalyze norm emergence and change, the major mechanism in the norm life cycle.

The catch is that while the true outcome is exactly the average of the predictions from the population, the outcome that each agent perceives is obscured by noise (a random draw from a uniform distribution bounded by zero and the specified maximum noise level). Noise can be thought of in two ways. First, it could be simulating a lack of information or uncertainty. Second, it could be conceived as representing the complexity of the social environment – the higher the noise levels, the less clear agents are on what the appropriate group outcome should be.

An additional aspect of the social context is the existence of a natural attractor or natural norm in this system. Rule 4, which produces predictions between 45 and 55, is a pre-ordained focal point or natural norm. (Some constructivists have argued that intrinsic characteristics of certain ideas make them more likely to become norms.) Averaging random numbers between 0 and 100 will produce a mean of around 50 in the long run, and thus agents should be drawn to this rule because of its intrinsic characteristics. The baseline model explores the conditions under which the agents can find this natural attractor through uncoordinated, adaptive behavior.

The real test of the logic of the norm life cycle begins as entrepreneurs are introduced into the model. (My model could easily be modified to examine other aspects of norm dynamics.) Norm entrepreneurs suggest a rule to the agents at specified intervals (every 50 rounds). Each agent replaces its currently worst performing rule with the norm entrepreneur's suggestion, and the suggested rule starts with a fresh score of 100. In the base version of the model, the entrepreneur is able to reach all agents simultaneously, and automatically convinces all the agents in the simulation to add the suggestion to their repertoire of rules. Crucially, the agents will only use the suggested rule if their other rules have been weakened through past punishments – just because a new idea about appropriate behavior is presented does not mean it will automatically influence behavior. From this baseline, I then also test the effects of

limited reach – how entrepreneurship works when entrepreneurs can only reach a portion of the population.

With the artificial agents and their environment created, simulations can be run to analyze how different parameter values influence norm dynamics. The model of the norm life cycle enables me to ‘experiment’ with various conditions that influence norm dynamics, such as the existence of a norm entrepreneur, the ability of the norm entrepreneur to make suggestions to the population, the levels of social complexity, or the sizes of populations, to name just a few possibilities. One advantage to this method, then, is the ability to fully control the parameters of interest in the artificial world and to create as many simulated histories as desired. Data creation is very simple when you can create 50 or 100 simulated histories that last 10,000 time steps in a matter of seconds. There is creativity involved here. Because one can create an infinite number of simulated histories, it is necessary to think carefully about which parameters are of most interest to test and setting up a protocol for producing simulations is crucial. The notion of knowing when to stop gathering data is analogous to the conundrum raised by Checkel, Dunn, and Leander (in this book) but is especially acute because the simulations can create as much or as little data as the researcher desires.

I used graphical analysis to picture how the group outcome changed over time. I used simple descriptive statistics to determine when norms emerged and how long they remained in existence. At a more sophisticated level, I also analyzed the data for the existence of statistical distributions called power laws, which reveal important features about the kind of process that facilitates norm emergence and change (Hoffmann 2003, 2005b). While space constraints limit a full reporting on the results of the simulation here, Appendix A presents some of the key modeling outcomes in visual form.

Beyond modeling norm dynamics

Modeling the norm life cycle is but one of an infinite number of possibilities for using ABM, because modelers have flexibility both to endow agents with decision-making procedures and to place them within an environment. ABM can accommodate *any* decision-making logic that can be formalized into algorithms. Agents can be almost entirely autonomous (following Post’s emphasis on psychology in this book), mostly calculated (along the lines of Checkel in this book) or significantly constrained by discourse and social context (consistent with most of the contributors in this book). For example, I have worked

with a land-use model that uses modified expected utility procedures and a spatially explicit, dynamic landscape (Hoffmann *et al.* 2002). Axelrod (1997) has worked with a model where the agents are essentially rational actors playing the prisoner's dilemma in a repeated fashion, in a soup. Lustick (2000) confines his agents to a grid and their decision logic is based on the local distribution of cultural expressions (represented by different colored agents). Cederman (1997), in his national emergence model, even allows agent boundaries to change as some agents subsume smaller agents. The *Journal of Artificial Societies and Social Simulation* provides a sense of the diversity of applications. The modeler's creativity and skill with programming are the only limits.

This flexible modeling of agency is what sets ABM apart from other computer-assisted approaches to social science that could be considered qualitative. For instance, systems dynamics models tend to neglect agency – there are no actors, as those models simulate stocks and flows of a system. The dialogic content analysis presented by Duffy (in this book) also uses algorithmic formalisms to study social life, but his analysis looks for patterns in empirical statements. Agency is represented by the content of speech acts, whereas ABM seeks to recreate agency artificially to simulate actors' behavior under a variety of conditions.

This discussion of the logic of ABM and a description of a particular model are more an introduction to the method than a defense of it. Although ABM has a number of advantages, there are a number of challenges that must be addressed, if not overcome. In the next section, I address these through the broader epistemological question which I frequently hear: Why on earth would you model constructivist IR thought?

Epistemology: why model?

It is entirely appropriate to ask modelers what one learns from their research. (Pepinsky 2005 asks this question explicitly.) ABM is *not* an empirical method, so the question is even more apt. The way I have described at least one type of ABM is that you start with an abstraction of reality (a verbal model like the norm life cycle), abstract it further to put it on the computer, and then simulate the second-order abstracted world. What kind of knowledge can such an exercise produce?

The ABM rubric can be considered a continuum between two ideal points. The first is what I call abstract heuristic modeling. With this approach, the models are kept very simple and no attempt is made

to capture empirical detail. In other words, neither the agents nor the artificial world are meant to represent specific real-world actors or situations. These models are used primarily as a way to explore fundamental logics. I wrote my norm life cycle model from this perspective (see also Axtell and Epstein 1994; Axelrod 1997; Cederman 2003). The other end of the spectrum is what I call flight simulator models. These are explicitly designed to capture important empirical features of actual histories and to create realistic future scenarios. For example, artificial agents can 'live' on GIS-informed landscapes (see Brown *et al.* 2005; Parker 2005). Given this range of possible models and approaches, I discuss three epistemological distinctions and a set of challenges associated with pursuing ABM in social research.

Heuristic/empiric distinction

ABM analysis provides a means for assessing the logic of verbal frameworks. We all approach empirical analysis with some abstract way of organizing information that tells a story about the world – a model. Most qualitative researchers use verbal models. For instance, the norm life cycle is a verbal model about how social norms emerge, diffuse, and evolve in a population. I used ABM as a way to assess its logic. Putting the essential elements of a verbal model into the computer helps to ascertain whether its conjectures are plausible and consistent, as parameters change in the simulations. This allows us to test a range of verbal models in a way that was heretofore unavailable.

Since ABM is not really an empirical research method, I envision its simulations as heuristic devices that provide insight into fundamental social processes. By simulating the norm life cycle, I am not learning about the emergence or evolution of any particular norm. Rather, I am learning things about the dynamics of norm emergence and evolution in general. ABM is perhaps best suited to testing explanations or verbal models in order to find 'candidate explanations' (Epstein 1999) for social phenomena. Such knowledge is crucial for empirical investigation. In this sense, ABM exercises can act as a plausibility probe or a pseudo-existence proof for our verbal models (Axelrod 1997; also see Klotz, in this book, on the role of plausibility probes in case selection). Also, ABM seems ideally suited to examining alternative processes, similar to the use of counter-factuals (as Checkel suggests, in this book). However, no matter how detailed the model, always remember that simulations are heuristics rather than true mirrors of reality.

Model/reality distinction

Despite the heuristic nature of the knowledge produced by ABM simulations, most modelers have a desire to learn something about the empirical world through this type of modeling. In other words, how do I know that I have a decent heuristic if not by comparing the model results from the artificial world to the empirical world? At some level, this implies at least some commitment to truth through correspondence with observation. ABM fits well within the tradition of scientific realism (Wendt 1999; Checkel in this book). But rather than just positing unobservable mechanisms, the modeling tool allows for a simulation of the unobservable mechanisms.

Axtell and Epstein (1994) have considered how to assess model results vis-à-vis the empirical world by thinking about how we can judge the worth of a candidate explanation. They offer four levels of model performance in relation to empirical observations:

- Level 0: Model generates agents whose behavior qualitatively matches 'real' agents being studied;
- Level 1: Model generates macro-patterns that qualitatively match empirical patterns;
- Level 2: Model generates macro-patterns that quantitatively match empirical patterns;
- Level 3: Model generates micro-patterns that quantitatively match empirical patterns.

Because I explicitly put a behavioral model onto the computer that has some basis in theorists' observations of empirical agents, this model reached Level 0 by definition.

My model results demonstrate Level 1 performance. The fact that emerging and evolving social norms are evident in the results gives us reason to be optimistic that the verbal model captures some important empirical dynamics. Note that this qualitative match means that the model produces results consistent with general empirical patterns, not specific empirical instances. To reiterate, the simulation exercises are not empirical tests. Rather, the question here is whether the model results track empirical patterns in any meaningful way. If the simulation model of the norm life cycle had not produced emerging and evolving social norms, this would have told us something important about the underlying verbal model (that perhaps it was not plausible).

I am not convinced that we can or should go beyond Level 1 for work with social phenomena. There are two pitfalls in trying to achieve quantitative correspondence with empirical events. First, it gives a false sense that you have gone beyond heuristics. The model is not reality and therefore the simulation will always be different from the empirical world, no matter how much detail is included. Attempting to 'tune' a model to empirical details is fraught with peril. The best we may be able to do is something akin to Cederman's (2003) analysis of the distribution of wars in his model and in the real world. He did not recreate World War II, but instead, his model produced a distribution of wars that corresponds with the distribution of wars in the historical record (size of war, not necessarily the sequencing of wars). The match of the distributions lends plausibility to his explanation for the dynamics of conflict.

Second, the more detail that is included, the harder it is to follow the dynamics of the model. It becomes more difficult to decipher which factors are driving the results as more parameters are added. Thus, I follow Axelrod's (1997) KISS principle: Keep It Simple, Stupid. Simpler, generic models capture fundamental dynamics. Deciding whether or not the dynamics observed in the model are 'real' or evident or plausible in actual social systems can only be done through empirical work. Any number of empirical methods discussed in this book (especially process tracing, ethnography, or pragmatic analysis), depending on the research question and phenomenon being simulated, can be used to ascertain this correspondence.

Indeed, the correspondence between a model and reality beyond Level 1 may not be an important question. The non-empirical nature of ABM for those doing heuristic (rather than flight simulator) modeling is not necessarily a problem. ABM provides a series of conjectures drawn from an artificial world that was constructed from a theoretical framework – an exercise in testing logical consistency and a search for fundamental dynamics. Given the absolute control that modelers have over the (initial) parameters of any model, it is possible to create multiple histories and to 'push' the logic of a verbal model in a variety of ways. This provides the social laboratory generally denied to social scientists (outside the narrow confines of some experimental economics).

This social laboratory can be enormously useful. By experimenting widely with the model parameters, it is possible to discover boundary conditions and novel empirical hypotheses that may be unavailable upon inspection of the verbal model. In the norm life cycle, for example, I can push the framework to find out how it works under different

conditions. I attempted to ascertain under what conditions entrepreneurs are necessary to catalyze norm emergence and change. The results of these simulation experiments can then be probed empirically.

Subject/object distinction

ABM confounds some of the normal lines of epistemological debate, notably the subject/object or researcher/researched distinction. The researcher in ABM simulations is god. In the models, I define everything about the artificial world. The outcomes may not be known in advance, but it is indisputable that they flow directly from the design of the artificial world. I very explicitly influence my subjects. There is no question or argument about whether the observer influences the observed – I created the observed. Yet there is a radical separation of researcher/researched, because my *observation* of this created world (as opposed to my initial-ization of the model) does not influence the outcomes. The artificial reality is objective in the sense that the model exists on the computer, where it is a closed system. During the course of most simulations, the researcher merely waits for the agents to run through their history. (It is possible to design models that are more interactive.) The agents I create do not respond to me as people or organizations being researched do.

This control is what makes flight simulator modeling attractive, if problematic. If one could capture essential elements of an actual social system, it would make testing policy scenarios relatively quick and easy. Applications have been developed for human-model interaction as well as scenario testing in a number of settings, giving one the sense of capturing reality. But given that the model can never capture reality in entirety (the simplest model of the world is the world), control of simulations can lead to hubris. Thus ABM is not a panacea for the methodological quandaries that face international relations scholars, nor is it suitable as a stand-alone method for all (or even most) research questions.

Potential limitations

A number of pitfalls with ABM must be at least considered. These range from the theoretical to the practical and professional. I will focus first on the potential reification of agents, then turn to the start up costs and the sociology of the discipline.

There are serious concerns about agent reification when representing social reality through ABM. Indeed, the focus of ABM is the decision-making of individual, *autonomous* agents, and Jeffrey Checkel and

Stefano Guzzini, among others, warn constructivists against implementing their insights through individualist behavioral models. Guzzini (2000: 150) cautions against 'mixing an intersubjective theory of knowledge with an individualist theory of action,' while Checkel laments that 'all too many constructivists rely' on behavioral models that 'are decidedly individualist in nature' (2001: 561). At first blush, these concerns seem to be damning for ABM. Its social structure is often a very simple aggregation of agent actions, something Guzzini blames for 'individualist reductionism' (2000: 164).

Yet ABM does not necessarily reify agents. They often change their preferences and behavior throughout a simulated history, and they can even change into other 'kinds' of agents. And while social aggregation is usually explicitly modeled in a simple fashion, it does not have to be. The simplicity of its representation of structure is, in some ways, less important than the fact that social structures are explicitly emergent; agency constitutes social structure. When agents' attributes are tied to the social context, simulation allows researchers to get at mutual constitution without bracketing. Neither the agents nor the structures are held constant.

In addition, focusing on internal decision rules does not necessarily equal an individualist ontology. Constructivists and others interested in language and meaning have struggled with how to characterize individual agency and the logic of appropriateness. Rationalist thinkers rely on methodological individualism, but when agents are socially constituted through intersubjective reality, the task becomes more difficult. Yet, these ontological differences may have been overstated. Even Kenneth Arrow claims that '...individual behavior is always mediated by social relations' (1994: 5). One change in the verb – from 'mediated' to 'constituted' – would make it amenable to constructivists. As Checkel concedes, 'where to draw the line between individual and social ontologies is no easy task' (2001: 559).

Beyond such theoretical concerns, it is imperative to get enough experience with modeling to pursue interesting tests of verbal models and to avoid 'programming in' the results. As it stands now, there is no way to begin ABM analysis without learning both a computer language and principles of programming agent-based models in at least a rudimentary way. Just like learning any language, one needs to grasp the vocabulary, grammar, and syntax of computer languages and learn how to make coherent statements (or algorithms). Unquestionably, this is a barrier to the widespread adoption of ABM; it requires a commitment similar to that of area specialists who must become at least functional in a foreign

language and immerse themselves in the life of their research area. (On investments in cultural competence, see Neumann in this book.)

Fortunately, the barrier is not insurmountable. All computer science departments offer basic courses, the ABM platforms are widely available as freeware on the web, and texts for all skill levels can be found in any major bookstore. Many of us who use the method do have prior engineering or natural science backgrounds. For instance, I earned a degree in environmental engineering. For those without programming experience, the time investment will likely be more than a year to get up and running.

In terms of the technical specifics, most ABM scholars use an object-oriented programming language like Visual Basic, C++, or JAVA. The object-oriented aspect is crucial because it allows the programming of varied, individual agents. Recently, a number of ABM programming platforms have proliferated (ASCAPE, SWARM, REPAST, MASON), allowing researchers to choose between building a model from scratch or taking advantage of infrastructure developed by others. The advantage of using such a platform is that many functions are already programmed; for instance, it is not necessary to create agents from scratch. The disadvantage is that the modeler loses some control over the design of the artificial world. I built my simulation of norms from scratch, using visual C++. To get a sense of what this language looks like, see Appendix B for a piece of my norm life cycle model code.

Finally, there are disciplinary challenges. The sociology of the discipline of international relations has created a series of (false) methodological dichotomies that have real consequences for researchers. ABM challenges these dichotomies and therefore offers a useful corrective. However, challenging the status quo is not always the safest course of action. While ABM has the potential to be a bridge between different methodological and theoretical inclinations, it can also be eyed with suspicion by both sides.

Conclusion

ABM is a technique that provides a rigorous alternative formal method for exploring theoretical insights. I find invaluable insights that I could not obtain by going directly from verbal models to empirical research. However, it is not a tool to be taken up lightly – I will not counsel that everyone learn to program computers. The objections to this type of modeling are not baseless and the modeling endeavor should be modest, but I find that the potential benefits of using ABM exceed the pitfalls.

Because ABM on its own is heuristic, it forces researchers who want to do empirical work to be methodologically diverse. I do not really see alternatives to modeling as much as I see complementary approaches. Empirical work is necessary in both the design and the analysis of the simulation models. Numerous methods can be used in conjunction with ABM simulations; the choice is a matter of research question. Indeed, the choice of tool should *always* be subordinated to the research question. For my work with social norm dynamics, process tracing is the natural complement. Tracing the norm life cycle empirically – picking a case or two and tracing what the verbal framework and modeling results say could/should happen – allows me to directly address the empirical research questions that drove the modeling in the first place.

This provides a recursive process. The modeling gives me a series of hypotheses or boundary conditions about generic norm dynamics to explore, and the case studies provide an empirical test of them. In turn, the empirical work provides new insights for further modeling experiments. Thus, we should not be thinking either formal analysis or case studies. Instead, the formal analysis enhances the empirical analysis, and in turn the empirical analysis should inform further modeling efforts. Combining the insights garnered with both methods provides the best analysis.

Acknowledgment

I thank Alice Ba, Owen Temby, Paul Hartzog, the editors, and members of Audie Klotz's 2005 graduate methods seminar for their insightful comments on earlier versions of this chapter.

APPENDIX A: MODELING RESULTS

Life without norm entrepreneurs

When norm entrepreneurs are absent from the system, two types of macro-patterns emerge in the simulations. Depending upon the noise levels in the system, the simulation exhibits a strict dichotomy between stability and volatility in the system. Figures A.1 and A.2 are typical runs without norm entrepreneurs. As the noise in the system increases, the simulation switches from stable to volatile. Each figure reports the average predictions (group outcome) made in each round by the agent population over 1000 rounds. Each of these simulations was run with 10 agents and a precision level of 5 per cent. The only variable altered

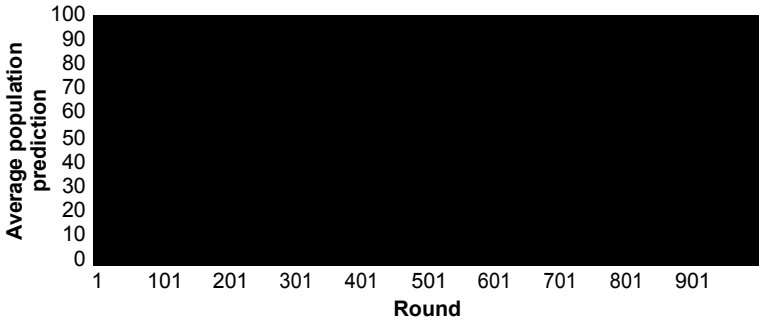


Figure A.1 Population predictions – low noise, no entrepreneur 10 agents, 6% noise.

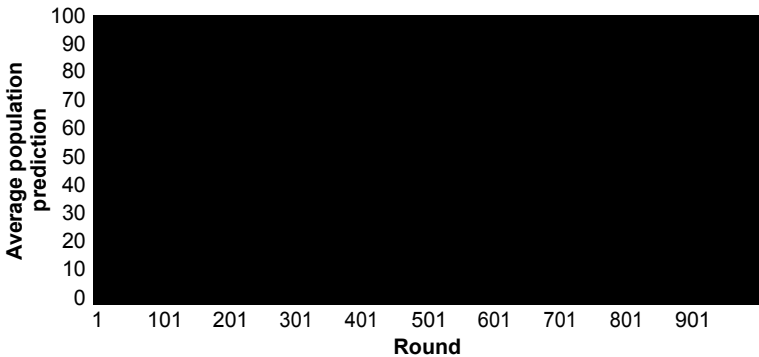


Figure A.2 Population predictions – high noise, no entrepreneur 10 agents, 10% noise.

from run to run was the level of noise added to the average prediction. In Figure A.1, the predictions reach a stable level relatively quickly as the agents arrive at the same rule. As the noise increases in Figure A.2, however, the agents are unable to come to agreement and thus the average predictions fluctuate wildly.

These figures demonstrate that when the noise level is low enough, the agents hit upon the dominant rule (often very quickly) in the system, rule 4. As the noise increases (as the agents are less able to see the true outcome and are thus less certain about its appropriateness), the agents are unable to come to agreement on any rule and the average prediction reflects this uncertainty. The agents are unable to find a rule that can be intersubjectively agreed upon and thus the agents continually cycle through rules.

Without norm entrepreneurs the agents' actions produce either a volatile or a incredibly stable macro-pattern with a strict break-point between the two types of patterns. The macro-patterns, in turn, alter/reinforce agent behavior and identity (constituting agents) leading to cycling in rule use or the domination of a single rule. The dynamism of the system is either out of control (volatility) or disappears (stability). We see the natural norm emerge or no norm at all. However, this simple set of outcomes does have interesting implications. The model suggests that there are situations where norm entrepreneurs are entirely *unnecessary* for norm emergence. When an idea is intrinsically attractive and the social complexity is low enough such that all the agents can appreciate the attractiveness of the idea, the idea can become a norm without any entrepreneurial effort.

Life with norm entrepreneurs

In contrast to the dichotomous patterns exhibited when the system lacks norm entrepreneurs, their presence creates different patterns. First, norm entrepreneurs are able to influence which rule rises to dominant status when the noise/precision levels would otherwise lead to stability around the dominant rule. See Figure A.3 for a demonstration of this effect. The simulation depicted in Figure A.3 is similarly configured to the simulation run in Figure A.2, except that a norm entrepreneur is now present.

The impact of the norm entrepreneur was significant. The agents still 'crystallized' around a single rule for the majority of the simulation, but

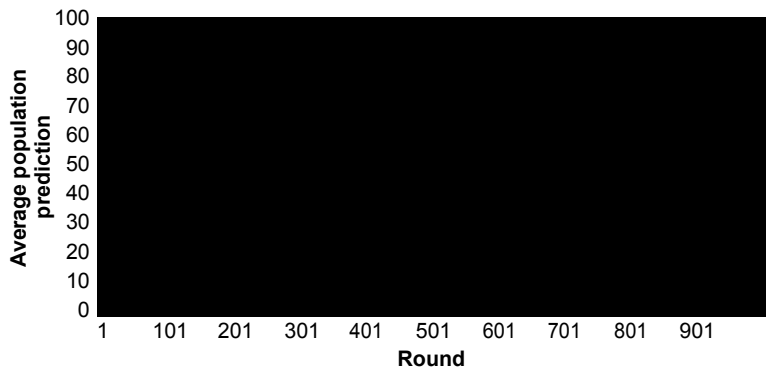


Figure A.3 Population predictions – low noise, no entrepreneur present 10 agents, 6% noise.

instead of the dominant rule 4, the agents crystallized around rule 2 (which returns a prediction between 15 and 25) after the suggestion of the norm entrepreneur. The norm entrepreneur was able to alter the manner in which the agent population crystallized around a single rule – a rule that generates a prediction different from the otherwise dominant prediction that hovers around 50. Repeated trials demonstrated that *any* of the rules can rise to normative status under these conditions.

Lock-in is not the only effect that norm entrepreneurs can have on the system. At higher levels of noise, entrepreneurs catalyze metastable patterns in contrast to a strict breakpoint between volatility and stability. Norm entrepreneurs allow the system to walk the line between volatility and stability and they create patterns of rising and falling norms over time. Metastable patterns occur when pockets of stability arise but do not last – there is stability in the system but it is not robust. In these simulations, the agents can coalesce around any of the rules and we see the rise and demise of intersubjective agreement among the agents. In essence, the norm entrepreneurs are able to catalyze intersubjective agreement, but the agreement does not ‘dampen’ the dynamism of the system. Instead, the agreement (or norm) lasts for a while before eroding via agent choices and new norm entrepreneur suggestions. The stability erodes because the system is too noisy to support long-term stability and norm entrepreneurs periodically prod the system with new suggestions. Norm entrepreneurs are thus able to catalyze both norm change and norm evolution.

Figure A.4 demonstrates the impact of norm entrepreneurs on a simulation similar to the one run in Figure A.2. Here a metastable pattern of

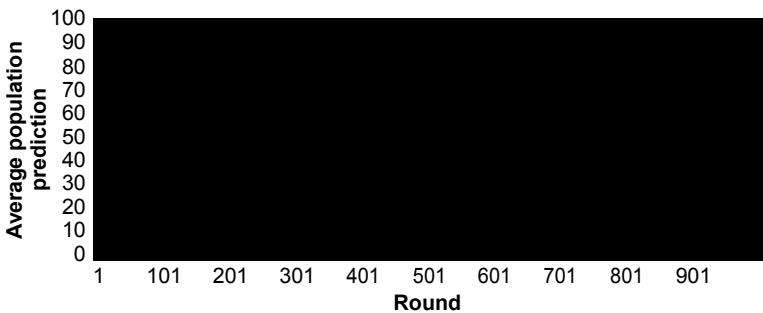


Figure A.4 Population predictions – high noise, no entrepreneur present 10 agents, 10% noise.

emerging and dissolving norms is evident instead of cycling, a pattern that results from the norm entrepreneur's suggestions at a level of noise high enough to cause volatile outcomes in systems lacking an entrepreneur. The norm entrepreneurs catalyze periods of intersubjective agreement among the agents – they make it possible for agents to crystallize around a rule for relatively short periods in an environment that would otherwise lead to volatile patterns.

APPENDIX B: NORM LIFE CYCLE MODEL CODE

This part of the code enables the agents to decide whether to take the entrepreneur's suggestion. The agent is given the suggestion and determines whether it will take the suggestion by comparing its current susceptibility to suggestions against its attribute for baseline susceptibility. If it takes the suggestion, the agent determines which rule to change. This is one of the functions of the agents and each agent runs through this function, though with different outcomes due to agents having different attributes. The entire code consists of about 500 lines (a relatively short model).

```
void agent::alter_rules(int suggestion)
{
    int change = 0;
    int choice;
    int susceptible;
    susceptible = rand()%100;

    if (susceptible <= susceptibility){
        if ((score[0] < score[1]) && (score[0] < score[2])){
            rules[0] = suggestion;
            score[0] = 100;
            change = 1;
        }
        if (change == 0){
            if ((score[1] < score[0]) && (score[1] < score[2])){
                rules[1] = suggestion;
                score[1] = 100;
                change = 1;
            }
        }
    }
```

(Continued)

```

    }
    if (change == 0) {
    if ((score[2] < score[0]) && (score[2] < score[1])){
        rules[2] = suggestion;
        score[2] = 100;
        change = 1;
    }
    }
    if (change == 0) {
    if ((score[2] == score[1]) && (score[1] == score[0])){
        choice = rand()%rules_avail;
        rules[choice] = suggestion;
        score[choice] = 100;
        change = 1;
    }
    }
    if (change == 0){
        if (score[0] == score[1]){
            choice = rand()%2;
            rules[choice] = suggestion;
            score[choice] = 100;
            change = 1;
        }
    }
    if (change == 0){
        if (score[1] == score[2]){
            choice = 1 + rand()%2;
            rules[choice] = suggestion;
            score[choice] = 100;
            change = 1;
        }
    }
    if (change == 0){
        if (score[0] == score[2]){
            choice = rand()%2;
            if (choice == 0){
                rules[0] = suggestion;
                score[0] = 100;
                change = 1;}

```

```
        if (choice==1){
            rules[1]= suggestion;
            score[1]= 100;
            change = 1;}
    }
    change = 0;
}
```

Part IV: Implications

This page intentionally left blank

13

‘Qualitative’ Methods?

Samuel Barkin

In this chapter, I neither present a method nor draw conclusions about the methods presented in the substantive chapters of this book. Rather, I reflect on the category around which this book is organized. The term ‘qualitative’ evokes a narrative or analytical richness, a method that brings out more detail and nuance from a case than can be found by reducing it to quantitative measures. But in practice, the term is generally used simply to mean ‘not quantitative,’ as Matthew Hoffmann notes in his discussion of agent-based modeling. Qualitative methods are, in this sense, a default category.

At first glance this categorization seems benign. What harm is there in a default category for methods that are not covered in the quantitative methods classes that so many graduate programs in political science require of their students? But the categorization is problematic, for two sets of reasons. The first of these is that to speak of qualitative methods is pedagogically counterproductive. It misleads students, and to the extent that we internalize the categorical distinction, it misleads researchers as well. The second set of reasons is that the phrase is politically fraught. To speak of qualitative methods is to stake a claim in the methodological disputes that divide the field of political science. Discussion of ‘qualitative methods’ becomes a proxy for claims about what does or does not constitute legitimate political science, because any method that fails to fit even into the default category cannot really be legitimate. To speak of some methods under the heading of qualitative implicitly but clearly stigmatizes others.

I should stress at this point that the argument here is about categorization – it is not about the legitimacy or utility of any particular method. Of course, a claim that discussion of ‘qualitative methods’ as a category inherently makes claims about what constitutes legitimate

political science by its own logic must make a claim about what constitutes legitimate political science. The perspective underlying this discussion is one of methodological pluralism, but at the same time of a need for methodological specificity. The conclusion will return to the idea of methodological specificity, making the argument that real pluralism is incompatible with the dichotomization implied by a quantitative/qualitative divide.

Pedagogy and qualitative methods

The use of the phrase 'qualitative methods' is often found in the context of pedagogy, of teaching people how to use method(s). And that is the primary point of this book. Therefore, an important step in assessing the effects of having this category (as opposed to whatever particular methods we may put in it) is to ask what the pedagogical effects are. Not only does this particular categorization not help in the teaching of international relations methods, but it can be actively misleading, for three general reasons. First, it implies that these methods have some core feature in common. This has the effect of highlighting similarities and obscuring differences. Second, it confuses the difference between analysis and research design. This has the effect of highlighting differences and obscuring similarities among methods that cross the quantitative/qualitative divide. (I'll comment below on the place of formal methods.) Third, it fetishizes method, which both contributes to the reification of particular methodological divides and privileges empirical analysis over theory.

There's no core

What does one teach in a qualitative methods course? Much the same as one puts in a book on qualitative methods – some of everything, except for quantitative methods. The range in this book goes from discourse analysis to personality profiling, from feminism to agent-based modeling. It is, after all, a default category. Let us leave for now the question of what gets left out of the course (or book) – I will return to that in the next section, on the politics of qualitative methods in international relations. Many approaches to the pedagogy of qualitative methods are self-consciously pluralist, and as such aim to include as broad a range of specific methods in the course (or book) as possible. (For an assortment of syllabi, see the website of the Consortium on Qualitative Research Methods at <http://www.asu.edu/clas/polisci/cqrm/syllabi.html>.)

Such inclusiveness, however, leaves one in a pedagogical quandary. One cannot cover all qualitative methodologies, if for no other reason than there is no discrete set of methodologies that one can claim to have covered comprehensively. At the same time, the more one strives for inclusiveness, the less one can do justice to most, if not all, of them, due to lack of time or expertise. That leaves a hodge podge that does not build on the sort of common core found in introductory quantitative methods classes. It implies, for example, that small-*n* inferential analysis bears more categorical similarity to Foucauldian genealogy than to statistical analysis.

The lessons are twofold. First, one is suggesting that there is a discrete number of qualitative methods that can sensibly be reviewed in the absence of a research question/focus. While it is true that courses in quantitative method also cover a variety of specific techniques, these build from a core that is taught at the beginning. Second, there is an implication that an understanding of individual qualitative methods takes only a week or two, unlike an understanding of quantitative methods, which takes a sequence of courses. So there cannot be as much to them.

The goal of a course on quantitative methods is clear: to teach a discrete set of techniques useful in analyzing certain types of data (those that have been quantified) once these data have been gathered. This raises the question of the analog for qualitative methods. There are no clear guidelines about how to interpret when using interpretive methods. Chapters in this book, ranging from Leander, Neumann, and Dunn to Checkel, ultimately rely on the good sense of the researcher, rather than clear replicable rules for deciding on issues of evidence and interpretation (although Duffy seeks to remedy this problem).

There are writing skills that are perhaps analogous to the statistical skills taught in data analysis classes. Learning how to write better may serve many of our students well, but that is not what we generally teach in qualitative methods classes. Or we might teach things like epistemology and research design that are not directly analogous to the quantitative techniques. For example, the categorical distinction between positivism and post-positivism would make much more sense if one went beyond a general survey. But questions of epistemology and research design are not best divided along quantitative–qualitative lines.

Analysis and research design

Since courses intended as qualitative equivalents of quantitative methods courses are, in a sense, inherently hollow, they tend to be filled

with things other than analytical technique. These range from data-gathering techniques, such as elite interviewing and content analysis, to discussions of epistemology and the philosophy of the social sciences. What these things have in common is that they are not inherent to the category of qualitative methods.

Teaching (and thinking about) them in the context of discussions of qualitative method, understood in categorical terms as the contradistinction to quantitative method, has the effect of confusing issues of analytic technique with broader issues of research design. Few issues of research design are specific to quantitative analysis, understood as the use of statistics, other than the need to find data that are quantifiable. If one understands quantitative analysis more broadly as international-relations-with-numbers (or more precisely with mathematical symbols), there are no issues of research design that are specific to it.

This confusion artificially delimits the flexibility of specific data-gathering techniques, in a relatively straightforward way. Few specific data-gathering techniques are suited only to statistical analysis. Typically, any information-gathering techniques can be used to generate either quantitative or qualitative data. Compare, for example, Hermann on content analysis and Duffy on pragmatic analysis (in this book). To use techniques only to generate quantifiable data would be to lose much, if not most, of the meaning and nuance in the information. Assigning the discussion of data-gathering techniques to courses on qualitative and/or quantitative analysis is either redundant (if done in both) or misleading (if done in one but not the other).

A second effect of the confusion of analytical techniques with research design is that it obscures distinctions in research design that do not correlate with a quantitative/qualitative distinction. As King, Keohane, and Verba argue in *Designing Social Inquiry* (1994), the requirements of research design necessary to substantiate inferential claims is the same whether or not the cases will be subjected to statistical tests. The need for care in the specification of variables, case selection, and data validity are the same either way. Checkel makes a similar case (in this book) on causal process tracing, although he notes differences between causal and correlational analysis that King, Keohane, and Verba fail to address. This is not to suggest that we should be focusing on research that makes inferential claims, only that many scholars of international relations do make such claims, and the requirements of research design to do so cross the quantitative/qualitative boundary.

Similarly, critical theory research looks not at 'objective' data, but at the discourses through which we understand the political; see, for

example, Neumann and Dunn (in this book). This sort of distinction also does not correlate with a quantitative/qualitative divide. Critical approaches are more reasonably introduced in general epistemology courses (and expanded upon in methods courses that focus specifically on the discursive), rather than lumped in qualitative methods courses along with small-*n* inferential studies with which they are epistemologically incompatible. In a discipline in which (particularly in the United States) quantitative methods courses are often required of graduate students but qualitative methods courses are not (Schwartz-Shea 2005), to discuss critical approaches primarily in the context of a course on qualitative methods is to allow students who focus on quantitative methods to avoid learning about it altogether.

The upshot of these observations is that there is an argument to be made for teaching epistemology and research design issues comparatively, rather than separately through distinct qualitative and quantitative methods courses. This still leaves scope for teaching particular techniques or approaches, be they analytical techniques such as statistics, data-gathering techniques such as surveys, or philosophical approaches such as critical theory. The common theme in all three of these examples is that they are organized around a core of ideas.

Fetishizing method

These three examples are not fully analogous, however. Statistics are a method of data analysis. Surveys are a method of data gathering. Critical theory, however, is not necessarily best understood as method. While it does involve method (which both Neumann and Dunn discuss in a gratifyingly accessible way), understanding critical theory also requires thinking about epistemology in a way that thinking about quantitative methods does not.

Statistical analysis can be approached from a mutually incompatible array of epistemologies, from logical positivism to philosophical realism, and a quantitative methods course can do a perfectly good job of training students in statistical techniques without addressing these epistemological distinctions. A course in critical theory cannot. This makes discussion of critical theory in a 'methods' course incomplete in a way that is not true of discussion of statistical techniques. A response to a prevalence of quantitative methods courses and literatures that focuses on qualitative methods as a category thus risks fetishizing method at the expense of broader issues of epistemology, methodology, and theory.

This is not to suggest that getting method right, and doing it well, is not important. But too great a focus on method can distract from other

key parts of the research process. For example, studying technique in the absence of a broader epistemological context can lead to a commitment to technique without a clear grasp of its uses and limits. Another part is theory and theory-building. Too great a focus on method biases our work toward empirical analysis and away from theory. More broadly, fetishizing method risks distracting us from the study of politics. One often sees statistically elegant studies of politically banal questions. Accepting the quantitative focus on method and mapping it onto non-quantitative approaches risks importing a norm that how you study international politics is more important than what you study. Beyond fetishizing method generally, the creation of 'qualitative methods' as a category in response to the prevalence of quantitative methods courses reifies that divide as the predominant feature of international relations pedagogy. This is problematic both because it is misleading, thereby leading to muddled thinking about epistemology and method, and because it is prone to becoming a focus of debate in the field, distracting from the actual study of international politics.

While the problems with qualitative methods as a category have been discussed above, quantitative methods as a category may seem more straightforward. Quantitative analysis is analysis of numerical (or quantified) data using statistical techniques. But this category is often used to refer to any approach that uses mathematical symbols. For example, game theory is often lumped in with statistical techniques, because both seem to be mathematically intensive, and practitioners of both are prone to claiming the mantle of science for their approach alone. This lumping is sometimes done on the qualitative side of the divide. Witness the absence of game theory in most qualitative methods courses (although not all – witness the inclusion of the Hoffmann chapter in this book). It is also done on the quantitative side. See, for example, the National Science Foundation funded Empirical Implications of Theoretical Models project, which is premised on the idea that good political science requires bridging the gap between formal modeling and statistical modeling, without addressing any of the epistemological issues raised by this premise (NSF 2002).

Other than a common use of mathematical symbols, these two approaches have little in common and are in important ways epistemologically mutually incompatible (MacDonald 2003). Lumping them together may make social and sociological sense, given the construction of the academic field of international relations in the United States at this point in time, but it makes neither methodological nor epistemological sense. The fetishization of method obscures these differences.

A reification of a quantitative/qualitative divide also leaves a number of approaches in a categorizational limbo (and missing from this book's attempt at inclusiveness). If we consider complex *game theory* to be a quantitative method, what do we do with narrative *game theory*? Is a discussion of the prisoners' dilemma quantitative or qualitative? And what about network analysis? At one level it should not matter at all – if one wants to do *agent-based modeling*, one should read Hoffmann's chapter, learn the requisite computer skills, and then just do it. But at another level, if we reify a methodological divide in the sociology of the field, fitting into neither category means not fitting into the field's map at all (as Hoffmann discusses).

This last observation leads to my second *general point*, and the next section. The categorization of methods is not just a pedagogical act. It is also a political act. It is not just about what *gets* put where, but about who *gets* put in which side of a dichotomy, and who *gets* excluded altogether. And these inclusions and exclusions affect who *gets* research resources, and who *gets* published.

Power and qualitative methods

Whatever the pedagogical effects of the creation of qualitative methods as a category, it is both a result and a cause of the politics of exclusion in the discipline of international relations. The creation of the category and its ancillary courses, books, and organizations is a response to the perceived privileged position of quantitative methods in various journals, academic departments, and funding organizations. It is also a cause of these politics, because discussion of what *gets* included in the category is in effect discussion of what constitutes real social science.

The Perestroika movement in Political Science is a case in point – it is a forum dedicated to the reform of the American Political Science Association, but at the same time it functions as a forum for the promotion of methodological pluralism against the dominance of quantitative methods in the discipline (Monroe 2005). The issue of reform suggests that the creation of the category of qualitative methods is, in part at least, a political attempt by those who do not use quantitative methods to improve their access to the professional resources of the discipline. (I discuss the question of pluralism below.)

To the extent that it is a political attempt, one can reasonably ask whether or not it is likely to be successful. The answer is unlikely to be an unqualified yes. Committing to a disciplinary politics of quantitative/qualitative divide has the effect of reifying a dichotomy between

scholars who use mathematic symbols in their research, and scholars who do not. As an exercise in political coalition-building, this is questionable. It puts all of those perceived to have privileged access to resources in one camp, thereby presumably reinforcing their incentive to cooperate among themselves to protect this privilege. It also cedes to them the mechanism for doing so, the mantle of 'science.' In a discipline in which claims to science are based on the sorts of symbols used, those same statisticians and game theorists are in a much better position to access the resources linked to the claim to science (the Empirical Implications of Theoretical Methods project comes to mind here). Helping to create this disciplinary geography is not necessarily an effective political move by scholars who do not use those symbols.

The reification of qualitative methods as a category not only helps to cement existing in-group/out-group dynamics within the discipline, it also creates tension within the out-group. It does this by defining the boundaries of the out-group. If a method for the study of international relations is neither quantitative nor qualitative, then by implication it is not really a legitimate social science method at all. As such, any attempt to define what constitutes qualitative methods is by implication an attempt to define away the legitimacy of any method not included. Since there is no core element to 'qualitative methods' as a category, discussions of the category need to enumerate methods. Inevitably some are left out. The process of enumeration thereby becomes a political process of defining the legitimate methodological boundaries of the discipline.

This process of exclusion is sometimes undertaken self-consciously. For example, in *Designing Social Inquiry*, King, Keohane, and Verba clearly claim that inferential logic is the only logic appropriate to the empirical study of political science, implying that non-inferential approaches are illegitimate. Similarly, works that associate 'qualitative methods' with interpretation (including, to a certain extent, this book) are in effect attempts to legitimate interpretive methods. But the process of exclusion can also operate by default, even when not intended. For example, a discussion of qualitative methods as a 'toolkit' of inductive research techniques has the effect of implying that theory-driven research, such as critical theory, does not involve actual method, and is therefore not really social science. The exclusion may be unintentional, but it has disciplinary political effects nonetheless.

The answer to this politics of exclusion is a politics of pluralism. Methodological pluralism is in a way the qualitative camp's response to the quantitative camp's claim of science. Yet it is ultimately a political

claim more than a methodological claim. Underlying a call for methodological pluralism is the idea that we should all have the freedom to do our research as we see fit, rather than the claim that the research that we do is all equally valid. Many of the methodological perspectives represented in the qualitative camp, broadly defined, are not themselves pluralist. Critical theory is no more sympathetic to behavioralist research, for example, than behavioralism is to critical theory. For that matter, even within the qualitative camp, critical and inferential methods are not mutually compatible in an epistemological sense (despite Klotz's attempt to reconcile them in the case selection chapter in this book).

In other words, the politics of pluralism in qualitative method is belied by the epistemology of pluralism in qualitative method. If one believes, following Robert Cox (1981), that social theory should be critical rather than problem-solving, the political call to pluralism generated by qualitative methods as category is selling the study of politics short in exchange for disciplinary gain. The benefits of methodological pluralism become an unexamined assumption rather than a question to be asked.

If not 'Qualitative,' then what?

My argument is not in favor or against any particular method or methodology. Nor is it in any way a critique of any of the chapters in this book. Method should be done well, and the contributors provide excellent guidance. My point is about categories. We should be cautious about investing too much in 'qualitative methods' as a category, because it can be pedagogically counterproductive, and it reinforces a disciplinary political divide that its adherents should be questioning rather than reifying.

But if not qualitative methods, then what? If I argue against the category, what is it that I favor? My answer lies in categories that are both broader and narrower. The broader ones are general '-ological' categories that do not assume particular divisions. And the narrower ones survey specific sets of analytic and research tools that have core foci upon which they build, rather than reviewing disparate tools that have little in common.

Our thinking about how to think about method should begin with principles of epistemology, methodology, and research design (somewhat like Part I in this book). In terms of epistemology, the major issues need to be thought about equally by scholars on both sides of the qualitative/quantitative divide. Similarly, many of the research design

issues apply across a variety of approaches and are equally applicable to research that uses qualitative or quantitative methods. Feminist scholars and quantitative scholars, for example, may use different terminology to discuss the need to make sure that information gathered is appropriate to arguments made, but there are similar research design issues either way. Thinking about and teaching these common themes helps scholars to communicate across the divide and to think about their work in a way that de-emphasizes the fetishization of technique.

Of course, some techniques do require much specific instruction. This includes statistics techniques and formal modeling, as well as interviews, participant observation, and reading documents in Chinese. And it includes critical theory. Furthermore, it includes some approaches that do not fit neatly into categories, like agent-based modeling. But there is no analytical equivalence among most of these techniques – they do not provide skills that are useful at equivalent stages of research. The narrow categories, then, should involve courses designed around specific techniques. Individual departments will not be able to provide courses in the whole array, but categorizing techniques as 'qualitative' or 'quantitative' will not change that. And losing the category of 'qualitative methods' need not eliminate comparative method, because that should be taught in the general '-ology' courses.

That the category of 'qualitative methods' makes some sense in a disciplinary sociology, as a response to a perception that 'quantitative methods' hold a privileged place, does not make it a good idea. Categorizations have implications, and the implications of this one are worth discussing before we reify it in our teaching as well as our research.

14

Practicing Pluralism

Deepa Prakash

The authors in this book have offered a variety of ‘tools’ for qualitative research. In this chapter, I offer ways to keep those tools sharpened. While my chapter, like the others, is informed by personal experience, I write as a ‘learner’ and as a representative of my peers. It is based on two integral elements of the process of writing this book: the qualitative methods course that was the testing ground for the manuscript and the students who took part in it.

Each chapter in this book was originally assigned in draft form in Audie Klotz’s qualitative methods course. The authors received feedback from not just one or two but nearly 40 graduate students. In turn, students used the chapters to probe their research questions. In both iterations of the course, they also had the opportunity to interact with practitioners of the various methods; in Fall 2005, this meant a chance to meet with the authors themselves. The entire process reflected the book’s emphasis on dispensing user-friendly advice. The ‘student voice’ was part of the book project from the outset, and it is only fitting that it is represented in the final book.

A note on the ‘methodology’ used for this chapter is pertinent here. I gathered my data (student responses to the course, the chapters and the category of ‘qualitative methods’) through a mix of methods. I first sent out an emailed questionnaire to all the participants of the course, followed by interviews with the people who responded to further probe their responses to the questionnaire. Finally, I held a focus group with 12 people to get in-depth feedback and to have a wider discussion about qualitative methods. The various illustrations and responses are presented anonymously throughout this chapter.

Let me also preface with some words about the representativeness of the group of students and their responses. The Political Science

Department at Syracuse leans toward the pluralism emphasized in this book and therefore is atypical of the discipline. The introductory qualitative methods course is required along with quantitative courses, and students are actively encouraged to take methods classes in other departments. However, in most other respects, the students 'testing' these chapters are broadly representative. They did not necessarily specialize in International Relations; there were Americanists and public policy specialists as well as students from other disciplines altogether. They were also at different stages of their graduate careers. Some had just entered graduate school and others were in the midst of writing dissertations. Most had never taken a qualitative methods class, while some had taken those offered by anthropology and sociology departments. Thus the teaching tools in this chapter are informed and inspired by the views of students encompassing a wide range of interests, experience and expertise.

In the first part of this chapter, I offer practical tools for learning and teaching qualitative methods, for students and teachers alike. These ideas are primarily based on the homework assigned in Klotz's course and the innovative ways that students applied them. I outline the assignments and pedagogical strategies that proved most helpful to students as well the frustrations they continue to face. In the second part, I engage in a wider discussion about the category of qualitative methods from the perspective of students. Responding to the concerns Samuel Barkin raises, I assess to what extent and in what ways the debates surrounding the quantitative–qualitative divide resonate with students.

Learning tools

Teaching and learning do not occur in the classroom alone and do not always come under the guidance of professors. As many authors in this book suggest, there is an inherent self-pedagogy. Few would say that they have learnt all they need about their respective methods, and many imply that learning methods is a lifelong process. In this sense, my initial foray into qualitative methods was auspicious.

The first time I explicitly thought about methods was as a first-year doctoral student getting ready to write a summer research proposal. I had a fair idea of the question I wanted to explore and the theoretical literature I wanted to engage, but 'methods' was a different problem altogether. I could summarize my method in one line: 'Read things, analyze them, talk about what they might mean.' I realized with dismay that this would not do. So I did what most grad students do when

confronted with academic dilemmas: I talked to my peers, scoured the Internet, and skimmed through primers on the subject.

Quickly I learned that I was going to do 'discourse analysis,' that there were various ways to do this, and that all of these involved a lot more Foucault than I knew! Three years later, with a qualitative methods course under my belt and having read more deeply, I am able to think about the methods in my research with some degree of clarity. But I remember how daunting it felt to go off and teach myself methods 'on the job' of writing a major research paper.

As I talked to my peers in preparation for this chapter, it became apparent that my experience was far from unique. Many of us begin our methods training in this *ad hoc* fashion, often while working on a large project, sometimes the dissertation itself. Like the authors in this book, we do (usually) manage to figure it out (more or less), but there *is* a better way. The following six 'learning tools' encompass our suggestions.

Start early

Taking an introductory class on specific methods – as part of thinking about research methodology generally – has obvious benefits. So the first lesson underscored by my peers is that it pays to think about methods early. Starting early gives you time to get the training necessary to do methods well. Just like learning a language, waiting until the third or fourth year places too much pressure on students. This is particularly true for the more 'cutting edge' methods such as ABM (Hoffmann), profiling (Post), content analysis (Hermann) and pragmatic analysis (Duffy), but applies to all methods.

While there is no cut off for learning methods, an early exposure to the basic terrain of qualitative methods enables students to make informed choices. In particular, it helps to dispel constraining preconceptions. For instance, studying the discourse of foreign policy elites about topic X by doing discourse analysis can be unduly limiting. One student had gone into the course convinced of discourse analysis as the appropriate method for his research but discovered that pragmatic analysis helped him clarify that he was essentially interested in the *dissemination* of discourse and language. Probing a research question from a variety of methodological approaches illuminates new facets of questions that would have otherwise remained unexplored.

Starting early is not as simple as it sounds. Many programs require students to take a set of quantitative methods classes as part of the course work, but do not have such requirements for courses on qualitative methods. Thus, students may take qualitative methods when they

are finally *able* to, making the decision to use a mix of methods less considered. So it may be worth it for students to lobby to have qualitative methods courses offered regularly and for these to be accorded equal status in programs.

Of course, the question of *how* early one should take methods courses is also important. Taking a methods course steeped heavily in epistemological debates may be intimidating and even counter-productive without knowledge of the basic vocabulary of the discipline. But as I shall argue below, a practice-oriented approach to teaching methods surmounts this problem to some extent.

Start small

Another tip from students is not to bite off more than you can chew. Hopefully, reading this book has enabled you get a sense of what QM entails and the range of techniques it contains. Maybe your interest has been piqued by one or more methods or maybe it has enabled you to rule out others altogether. However, you might not be ready to commit to any particular method yet, and maybe you are not sure how to proceed in the stage between learning methods and what to use in a dissertation. What do you do next?

One way is to apply these tools to a small project or a sliver of your larger project. Exposing your research question to the gamut of methodological approaches can be overwhelming if you are learning methods for the express purpose of writing a major research project, with a lot of complex data or a huge fieldwork component. This is where a homework-based approach may be useful. As students taking the Syracuse qualitative methods course reiterated, focused assignments were a way to get a taste of doing ‘real’ research within manageable parameters. The assignments in Klotz’s class (see the section on ‘Teaching Tools’) typically required students to do a small piece of research, analyze or make inferences from it, and then reflect on the limitations of the method for their research. The emphasis was on bite-sized pieces of research, making use of the resources and opportunities immediately available.

Another way to keep the proportions manageable is to rely initially on a few major guidelines for the method, rather than plunging into its entire history and philosophical underpinnings. For readers using this book outside a formal course, it may be useful to look up a few QM syllabi and identify one or two important texts that are typically assigned in such courses to read in addition to the method-specific chapters in this book.

Look around you

As Ackerly's example of how she 'teaches' curb cutting illustrates, with a little imagination, our every day environment can provide resources in learn new skills. Campuses are a great place to practice interviewing, discourse analysis or ethnography in arenas that approximate your research focus. For instance, one creative student interested in fair trade movements spent a few caffeine-infused days at the local university coffee shop talking to people about fair trade coffee, embodying a mini-version of the 'deep hanging out' Gusterson advocates.

For another student examining the discursive formations from conflict situations involving extremist groups, a sports bar became an interesting site in which to conduct his 'fieldwork' for the homework on ethnography. Since he could not feasibly hang out with a militant group, he 'thought of other situations in which people develop negative attitudes toward strangers based on certain qualities found in those strangers.' Luckily, the baseball playoffs were going on at the time, and the Yankees and Red Sox happened to be playing one night. As baseball fans will attest, it was an apt choice for his research!

Similarly, since flying to Washington DC to hang out for a day at the Department of State or studying the plethora of speeches, official documents and treaties that make up US foreign policy was not feasible, a student interested in the discourse of US foreign policy used the environs of the Maxwell School as an intriguing substitute. Simply listening more closely to the 'discourse' on International Relations and Foreign Policy in his various classes for a week enabled him to 'scrutinize the power structures and implicit norms' underlying it. So the third tip in teaching yourself methods is to practice and to use the resources around you, enabling you to save travel and funding for your large project.

Assess yourself

For students who have a less concrete idea about the intricacies of their interest areas, using resources around campus in small but innovative ways is a useful way to practice tools without investing too much time learning about substantive issues. Those who already had a good idea of the data they would be working with approached the assignments in more traditional ways. Here too, doing exercises proved revealing.

For those who already had some training in qualitative methods (say, through a sociology or education course), the exercises were a way to hone their skills and try out new tools. As one student who was working on his dissertation proposal put it, 'The innovative element was trying

to weave one large project across the homeworks enabling me to see my topic in new ways.' For others, the assignments opened up tantalizing avenues for research. One student was excited by the potential that agent-based modeling offered to analyze the patterns and effects of state interactions. Learning more about its techniques also exposed him to new theoretical literatures. Yet another student found *all* the exercises extremely frustrating, forcing her to reframe her research topic in a more accessible manner and to contend with nagging questions about the feasibility of her entire project.

Doing actual research is very different from thinking about it in your head. As the students found out, what looked simple or difficult in their heads was not necessarily so in practice. You may find out that you are not very good at something you thought would be easy to do. Maybe 'deep hanging out' with people you are normatively opposed to is harder than you may have thought when contemplating research abstractly, or it might prove transformative, as Gusterson found with nuclear scientists. If you are interested in Political Personality Profiling but discover that you are not good at reading people or lack the 'psychological mindedness' that Post notes is required, you may want to supplement with – or shift completely to – other tools.

The point is that the exercises can go either way; they can lead you to reject certain methods in answering some questions or they can prove to be successful from the very start. In all three cases, students learnt something about the limitations of their topic and of themselves that they would not have known without actively trying their hand. Exercises can give you valuable lessons about what your innate strengths and weaknesses are.

Read exemplars

Most of the students told me that the chapters in this book provided a road map for research, to be augmented by other works and illustrations. Once you get a sense of the basics and you decide to explore one or more of the techniques, a good next step is to read other works by the chapter's author, to get a sense of what such work looks like in detail. Another step is to read the works the authors cite as good examples of the use of their methods.

If you continue to find the method useful and want to learn more, it is then helpful to read some of the philosophical underpinnings of the works and to immerse yourself in, say, Foucault or Bourdieu. Even without taking a QM class, one can be attentive to methods in the books and articles that get assigned in topics courses. Read appendices and

methodological sections of works you find persuasive. These are simple ways to adopt tools that appeal to you and to look out for pitfalls.

Triangulate advice

Although most students identify certain 'go to' persons (such as advisors or methods experts in the department) to talk to about methods, there are other sources they can and should consult. Looking for additional resources and opportunities also benefits researchers by putting them in contact with peers wrestling with similar questions and, especially useful for graduate students, creating networks with established experts in their fields.

Like my experience with 'teaching myself' discourse analysis revealed, a vital source of advice comes from peers. For example, while talking with the focus group for this book, one student asked us where she should take someone she was going to interview for a research paper. Since the interviewee was a refugee and they would be talking about sensitive issues, she was worried about meeting with him in a crowded or intimidating place. A range of opinions and ideas offered by peers helped her figure out her concerns and how to address them.

Particularly for those of you intrigued by the 'boundary crossing' techniques, such as agent-based modeling or pragmatic analysis, get in touch with the scholars working with such methods. Typically, these are small, growing communities whose researchers are, as some of the students I spoke to found out, particularly willing to help students interested in their tools and to suggest future avenues.

Since most departments realistically cannot offer a full array of methods courses, many students I spoke with also recommended taking methods courses in departments such as Sociology, Anthropology, Women's Studies and Education. Thus if you were interested in Ackerly's chapter, for instance, find whether another department offers a course specifically on feminist epistemology or methods (if yours does not). In this way, researchers meet each other, enabling them to form working groups with people well versed with qualitative techniques of various kinds. Such inter-disciplinary study groups have proven to be invaluable to students for both the fresh perspective they provide and the exposure to new works with methodological insights. Interdisciplinary perspectives are particularly useful given the complex nature of the international relations.

The growing number of methods workshops at major conferences is another source of advice. Forums such as the 'methods café' at the International Studies Association and the American Political Science

Association's Qualitative Methods working group, along with a growing number of similar workshops at regional conferences, are great places to meet practitioners of particular methods and network with people that are working in similar areas.

And as others in this book have already highlighted, the Consortium for Qualitative Research Methods (CQRM), based at Arizona State University, is a venue for students to receive focused feedback on their projects and to network. Its website provides even more resources, including syllabi (with other ideas for homework and other exemplars from additional literatures). It is useful to ask whether your department is a member of the consortium and even to lobby for this if it is not, as this guarantees the departments some spaces at the annual training workshop.

Teaching tools

As the above section has shown, there are a number of ways to keep your methods tool kit sharpened. However, much training does happen in classrooms. So it is not surprising that students also have strong opinions on the kind of pedagogical strategies that work. In this section, I will outline the aspects of the course we took that students benefited from and identify continuing challenges, in order to help others designing their own courses. The bottom line: have more practice and less epistemological debate. This translates not just to the course content but also to the sorts of assignments that students do.

Syllabi

The vast majority of qualitative methods courses (at least in Political Science and IR) follow a certain pattern: a significant amount of time is devoted to the enduring debates about the philosophy of science. Students typically learn about these debates by reading secondary literature and exemplars of the methods. As Barkin points out, courses also tend to treat research design as the main goal of learning qualitative methods. Finally, most syllabi stick to what are presented here as the 'classical techniques,' thus reinforcing the qualitative–quantitative divide. Barkin's point that few courses teach students concrete skills or techniques, analogous to quantitative methods courses, is also borne out. (I base these claims on a perusal of syllabi on the CQRM website, <http://www.asu.edu/clas/polisci/cqrm/syllabi.html>.)

But this is not necessarily inevitable or desirable. There are ways to ensure that students learn how to 'do' qualitative methods, rather

than just think about them, and from my conversations with students, it appears that students would prefer this latter equation. Achieving the balance between epistemological debates and practicing tools is difficult and will be somewhat specific to each department. To some extent, the ability to practice reflects the structure of our department at Syracuse, because there is a separate foundational philosophy of social science course where students first encounter the broader epistemological debates. However, most departments have such a course. For those that do not, the qualitative methods course can become the default primary venue for these discussions.

It is important to note that the students I talked to were divided about how much of each they wanted. What is clear, though, is that they definitely want opportunities to try out the techniques. Their responses suggested that students probe the epistemological implications of the methods, even in a course that is practice oriented. This reflection comes as a natural part of the process of doing research. As one student put it,

Perhaps the biggest gain for me was an increased ability to think self-consciously and reflexively about the epistemological implications of these methodologies. How do I conceptualize language and communication? What are the tradeoffs involved in adopting a more positivist versus a more interpretivist research perspective? What kinds of textual elements – if any – can be ‘counted’ in a meaningful way?

Assignments

Qualitative methods courses typically make students undertake three types of exercises and assignments: book reviews and critiques with a focus on methods and research design; practice-oriented exercises; and a final research design paper. Students I talked to uniformly found the practice-oriented assignments most useful. As one student put it, it gave her the ‘first taste of field work’ and allowed her to *be* a researcher instead of acting like one. I should note that the course we took did not ask us to do a book review or critique, so I cannot comment on whether students might also find that useful. (Note, however, the advice above to look at exemplars.)

In Klotz’s course, the homeworks changed between the two iterations of the class, reflecting an increasing emphasis on practice. While the first iteration of the course more closely incorporated the epistemological debates typical of qualitative methods courses, the second iteration was more practice oriented. The emphasis in the short assignments was on

describing what was actually done, how students conceptualized certain terms, who they talked to, what inferences were made – the focus was on the details.

For example, the assignment for ethnography in Fall 2005 asked students to answer, first, ‘Which aspects of their projects initially seem suited to the tools of analysis generally associated with ethnography? Are other aspects less suited? Why – on what basis are you making these initial assessments?’ Second, the students had to ‘apply one technique in ethnography in some way that is relevant to their research.’ Students had to follow the guidelines from the readings to perform this exercise and write up ‘what they did, whether it worked well or not, and how they might expand and/or modify your application of the technique for your research project.’

The final section of the assignment then asked an epistemological question: ‘Ethnography is often associated with an epistemological position that rejects the notion that social phenomena, especially meanings, are stable enough to be categorized as variables. Based on your reading and your preliminary foray into the application of this “interpretive” method, what is your initial position on this debate?’ By the second iteration in Fall 2006, this third part had been pared down almost entirely and the emphasis was on how to do a small slice of ethnography:

For most of you, doing a little bit of participant-observation in an appropriate field-setting will not be possible for the purposes of this assignment. Instead, simply practice in an alternative setting: go to an organization’s meeting, hang out in Eggers café or the TA bays, or talk with your housemates. Try to find a setting that might enable you to probe a question similar to one that you might have in your research. Tell me what you did, what you hoped to learn, whether it worked well or not, and how you might expand and/or modify your application of the technique for your research project. Alternatively, you might conclude that ethnography doesn’t look like a viable technique for your research project; if so, tell me why.

Doing these small applications of ethnography meant that straight away we realized some of its strengths and limits for our research. For instance, for a few of us working on extremist groups and political violence, ethnography had clear limitations. Who could one interview and ‘hang out’ with when studying the discourse of Al-Qaeda? Should we rule out participant observation altogether? It forced us to make choices about how far we were willing to go and anticipate issues for any larger project.

The exercises also prepared us for one of the most memorable aspects for students: the chance to meet with practitioners of the various methods. In the Fall 2005 class, students met most of the book's contributors themselves, getting even more of a personal insight into the life of a researcher than they got from the remarkably candid chapters. In Fall 2006, the experts came from around the various departments in the Maxwell School. Drawing local expertise is a rich resource that most courses can easily capitalize on, yet few appear to do so (from what I saw in posted syllabi).

Such interactions with experienced scholars humanize and demystify the process of research. Instead of seeing a perfect finished product such as a book, students learn how seasoned researchers cope with problems and often make arbitrary decisions. Homework assignments, furthermore, were timed so that students had already attempted to apply a particular tool *before* the expert's presence in the classroom. This allowed for more sophisticated discussion and more precise questions.

As a pedagogical tool, this one seems like something that is relatively simple to do and that goes a long way in fostering the kind of dialogue on methodological pluralism envisioned in this book. It also helped us think about some of the issues involved in turning an initial small slice of research into a full project, as we were then asked to do in our final assignment, the Research Design Paper.

The standard final paper in qualitative methods classes often takes an 'as if' approach. The brief is, Pretend *as if* you have resources, funding, language capability, the ability to travel and enough time. How then would you design your research? While this has its uses, it can mean that students write idealized designs based more on reviews of exemplars than their own knowledge of the feasibility of their methods.

In contrast, the final assignment in Klotz's course focused on methods rather than a comprehensive proposal. Students were asked to write a terse grant proposal but only write up the methods section of it, selecting two of the techniques surveyed in the course. Students at various stages in their graduate career found this useful because they did not have to spend too much time on doing an exhaustive literature review or developing theoretical sections, which can be particularly difficult for students just beginning to think about their research topics. For teachers, such an approach ensures that they are able to give feedback on methods, rather than contending with a wide variety of substantive issues about which they have varying degrees of expertise.

However, there are also some limitations that result from this narrow focus on methods. Barkin raises the point that qualitative methods

courses are ill equipped to teach analysis or inference in the same way that quantitative methods classes can. The experience of my peers reiterates this concern. Reacting to the chapters as well as the QM course, students consistently brought up the problem of not knowing what to do with data after they had collected it. Their feedback was that while they learnt how to conceptualize and collect data, they were less certain about the analysis part. Are there rules for analysis that correspond to the way statistical data can be read? How does one standardize analysis of qualitative data? Is this even possible or desirable?

Most qualitative methods classes do not focus on these questions in a general manner because they remain (understandably) geared to each student's research interest. Assignments ask students to apply methods like historiography or process tracing to 'some aspect relevant' to the students own research. For the most part this approach is necessary because students need the methods for their own work and are unlikely to be motivated by working on some general data set or to encounter the questions pertinent to their own topics in generic data. However, this means that students cannot see whether the frustrations they experience in analyzing and understanding data through a particular method are unique or general problems.

One solution is to have more of a mixture of assignments, whether in class or as homework, with most geared toward the student's own research but one or two uniform to the entire class. Possible assignments could be for all students to read through the same textual sources with the aim of producing a short analysis or to work in groups on one of their peer's research topics. Exercises such as coming up with a few words to make a dictionary from a commonly assigned text as part of the content analysis section shows to what degree methods can be replicable or not. Such exercises demonstrate the inherent variability of interpretation and analysis as well as opening a discussion on how to assess the validity and persuasiveness of interpretation, and whether this is possible at all. These collective exercises can also be done in class, achieving the balance between individual and group assignments. Another more commonly used strategy is to have students critique methodological sections of already published works. Assigning one book review from a choice of three or four books gets at the question of assessing method and its analysis.

Finally, it is possible simply to address the 'what do I do with this data' question explicitly in class. Perhaps the rules of analysis or inference do not directly translate from statistical methods. If so, what are the ways that students can assess the validity of their interpretations? What are

the uncertainties inherent in this approach and how have scholars dealt with them? A discussion on this topic would go a long way in addressing this concern of students and allaying their fears that there is a correct way to analyze data.

Such a discussion may actually lead to the answer that it is not necessarily bad to come away with questions about how to analyze data. These are precisely the sorts of questions that can be probed with advisors, in topics courses, and as you read further into the literatures relevant to you. It is misleading to think that one course or one phase in graduate school can tell you all you need to know about analysis. Instead, conceiving of research as a multi-stage process alleviates some of these anxieties.

Advanced courses

Teaching qualitative methods is difficult, as Barkin points out. How does one do justice to methods as rich and varied as discourse analysis or ethnography in one semester? Yet it was resoundingly clear that students want courses on qualitative methods, even if they do tackle ethnography in one week. They also want more than just a basic course and are acutely aware of the need for more training, thus echoing Barkin's concern that courses become a 'hodge podge' of rushed sessions. These concerns may be addressed by offering *advanced* qualitative methods courses. A couple of syllabi are available on the CQRM website, but for the most part these seem rare and mainly focused on comparative case studies.

Students I spoke with had some ideas for what an advanced qualitative methods course could look like. It might involve a semester long course focused on an individual method, such as discourse analysis or ethnography. Given the constrained resources of most departments and the unlikely scenario of having resident experts in all techniques, one option may be to team-teach courses structured in month-long modules for each method. (Professors and departments would have to make arrangements for teaching credit.) Another option would be to divide students into groups by methods and give readings and assignments to these groups within the class, along with some common topics that the class can do together.

'Qualitative' methods as category: 'the view from below'

Students are acutely aware of the power stakes in qualitative and quantitative methods. My peers recognize that being seen as 'qualitative researchers' puts us in a minority in American academia (though perhaps

less so elsewhere), with disadvantages on the job market and for being published in many mainstream journals. And we agree with Barkin that categorization is a political act.

But this does not mean that we are unhappy with the category. It helps us compete and cope with the demands of the profession. And it helps us avoid working in an isolated bubble. Students I talked with saw the category as giving them a voice and a vocabulary in academia. Many expressed a sense of relief at being able to find language that allows them to justify their choices to an outside audience.

In concrete terms, training in 'qualitative methods' gives us the vocabulary to do things like apply for grants (and perhaps attend a major annual workshop). A student from the history department told me that he received more grants than his colleagues, who receive little methods training, after being able to explicate what he was doing. This training also helps us to get proposals passed by committee members who may question the legitimacy of discourse analysis or to give job talks confidently to a potentially critical audience. Having a category that is intelligible to the mainstream, no matter how nebulous or flawed it may be, is better than a situation where quantitative and formal methods are juxtaposed with 'the rest' (especially given the artificial nature of these boundaries, as Hoffmann argues).

So while we are aware of the pedagogical and political implications of categorization, it has its advantages. But this is not the only reason for our support of the qualitative methods category. We recognize that exposure to different tools genuinely allows us to explore our questions and illuminates new ones. Students I spoke with unanimously believe that qualitative methods should be a required course in the same way as quantitative courses are, because both types of methods make us better scholars. And that means we believe that students who identify themselves as quantitative researchers would also benefit from exploring their questions with different tools.

Barkin's point that qualitative methods courses give the impression that one can understand individual qualitative methods in only a week or two is indeed troubling. One of the more awkward silences during the focus group was when I asked my peers whether they thought they could do qualitative research before they had taken a course. Quite a few of us agreed that we came in with the sense that you can do qualitative methods to some extent without training. Only after we took a course did we realize that we were wrong; you are not doing qualitative methods if you can read a book and interpret it. Barkin should be reassured to know that as students, the introductory course demonstrated how much

more training we need. It did not give us the (false) impression that we were now well versed as qualitative researchers.

Nevertheless, I was struck by a tension in how my peers do view qualitative methods as a residual category. Particularly for those of us who identified ourselves as users of qualitative methods, there was a measure of defensiveness coupled with a sense of superiority. Students can be quick to decry quantitative work as being shallow and 'easy' but powerful and 'legitimate,' while they see qualitative work as more esoteric and deep, but undervalued. The politics of categorization is not benign, and it will not allow students to truly embrace pluralism until the power disparities are reduced. Thus the lack of conversations between researchers vested in particular methods noted in the Introduction and by Barkin is perpetuated by students too, making the challenge to overcome these barriers and embrace genuine methodological pluralism even harder.

I remain more optimistic than Barkin, however, because most students enter graduate school with few preconceived notions or rigidly held beliefs about methods. It strikes me as ironic when students are told to be open to mixed methods, not to be too dogmatic, and not to fetishize. In the student view, these are hallmarks of the academic environment into which we are slowly socialized. Few of us come in with firm positions on epistemological debates; instead the politics of categorization and the necessity of picking a camp happen as we seek a place in the profession. As students, we are the most open to pluralism as we enter graduate school. If pluralism is the goal, our initial training needs to reinforce dialogue across methodologies.

Conclusion

In keeping with the candid and self-reflective tone of this book, let me conclude by recounting my own story of the methods I employed for this chapter, in a far messier manner than the contributors to this book would prescribe. As Leander might have predicted, my plan was ambitious. I was optimistic about the responses I would be able to collect as I sent out an initial questionnaire to all the students who took the class in the two years it was offered. And I had mental plans to expand my research to other universities. My initial optimism was dampened when less than half the people who received the questionnaire wrote back, despite several pleas.

Here is where my 'cultural competence' started to come into play; fortunately, I had learned a trick or two from Neumann and Gusterson.

As a grad student I knew that expecting people to send detailed responses to my questions toward the end of the semester was unreasonable. Interviews with people during their breaks were one way to address the problem. Finally, I decided that a focus group approach might be the optimal strategy, given the constraints of all the participants, as well as my own goal of getting a range of feedback on teaching tools.

By now, my expectations for people's participation were much more realistic, reflecting an adaptation based on my preliminary experiences in my 'field' (graduate students). I resorted to a combination of cajoling, nagging and offering material incentives (donuts and coffee, the grad student's ultimate weakness). Finally, a dozen people sat down with me and talked about their experiences. Despite my best attempts to survey a wide range of opinions, my 'sample' or focus group really consisted of my loyal friends.

With more funding and resources, I may have been able to garner more participation, perhaps from graduate students from other universities or even countries. I was constrained by money, time and my other commitments. Yet, the focus group session produced a free-wheeling discussion about the adequacies of qualitative methods as a category, what worked in the course, the use of the homework assignments, and the other comments I have provided in this chapter. The insights gained through this 'triangulation' of methods – questionnaire, interviews and focus group – were pivotal, confirming one of Checkel's main points. All in all, my partial survey was infinitely better than no survey at all!

With this illustration in mind, let me offer two final thoughts that are easy to forget in the concern with achieving rigor in methodology. My first point may seem strangely placed at the end of a book on (qualitative) methods: We must remind ourselves not to get too obsessed with methods. Students are aware of arguments, such as Barkin's, pointing to the dangers of method driven research. Often this argument is associated with quantitative research, but the same argument also applies to 'qualitative' methods. What initially excites us as researchers and makes us pursue grad school are questions that bedevil or anger us, regions that fascinate us, and puzzles that remain unresolved. Treat methods as tools to put the pieces together, not as the puzzle itself.

Second, it is worth keeping in mind that no matter how explicit the recommendations of the authors in this book and other works are, these are recommendations and not blue prints or formulas. They will work differently for you. As one student found while conducting a survey for the World Values database, 'cultural competence' means knowing when to break the rules. Conducting surveys in her native Turkey,

she found, was nothing like the formal protocol handed to volunteers, which forbade surveyors to enter participants' homes or to engage in prolonged conversations with them. Instead, she found herself participating in ladies afternoon gossip sessions and drinking tea in various homes, all the while learning much more than the survey could have ever revealed. Apart from leaving her skeptical about the 'objectivity' of surveys, her experience reinforced the lesson that research requires flexibility, humor, persistence and a little bit of a *laissez faire* attitude.

Acknowledgment

I thank all my peers who helped me write this in many different ways, particularly my focus group: Hannah Allerdice, Angela Fitzpatrick, Seth Fischer, Matt Guardino, Mike Makara, Ryan McKean, George Mitchell, Paloma Raggo, Eric Rittinger, Braden Smith and Matthew Smith. Special thanks to Asli Ilgit and to Audie Klotz for invaluable suggestions, ideas and chats.

Bibliography

- Ackerly, B. (2000). *Political Theory and Feminist Social Criticism*. Cambridge: Cambridge University Press.
- Ackerly, B. (2001). 'Women's Human Rights Activists as Cross-Cultural Theorists.' *International Feminist Journal of Politics* 3: 311–46.
- Ackerly, B. (2007a). 'How Does Change Happen? Deliberation and Difficulty.' *Hypatia* 22: 46–63.
- Ackerly, B. (2007b). 'Sustainable Networking: Collaboration for Women's Human Rights Activists, Scholars, and Donors.' In *Sustainable Feminisms: Enacting Theories, Envisioning Action*, ed. S. Sarker. Oxford: Advances in Gender Research, Elsevier.
- Ackerly, B. (forthcoming 2008). *Universal Human Rights in a World of Difference*. Cambridge: Cambridge University Press.
- Ackerly, B. and J. True (2006). 'Studying the Struggles and Wishes of the Age: Feminist Theoretical Methodology and Feminist Theoretical Methods.' In *Feminist Methodologies for International Relations*, eds B. Ackerly, M. Stern and J. True. Cambridge: Cambridge University Press.
- Ackerly, B. and J. True (forthcoming). *Doing Feminist Research in Social and Political Science*. Basingstoke: Palgrave.
- Ackerly, B., M. Stern, and J. True, eds (2006). *Feminist Methodologies for International Relations*. Cambridge: Cambridge University Press.
- Adcock, R. and D. Collier (2001). 'Measurement Validity: A Shared Standard for Qualitative and Quantitative Research.' *American Political Science Review* 95: 529–46.
- Adler, E. (1997). 'Seizing the Middle Ground: Constructivism in World Politics.' *European Journal of International Relations* 3: 319–63.
- Adler, E. (2002). 'Constructivism and International Relations.' In *Handbook of International Relations*, eds W. Carlsnaes, T. Risse, and B. Simmons. London: Sage.
- Agathangelou, A. (2004). *The Global Political Economy of Sex: Desire, Violence, and Insecurity in Mediterranean Nation States*. New York: Palgrave Macmillan.
- Alderson, K. (2001). 'Making Sense of State Socialization.' *Review of International Studies* 27: 415–33.
- Alker, H. (1988). 'The Dialectical Logic of Thucydides' Melian Dialogue.' *American Political Science Review* 82: 805–20.
- Arrow, K. (1994). 'Methodological Individualism and Social Knowledge.' *American Economic Review* 84: 1–9.
- Arthur, B., S. Durlauf, and D. Lane (1997). *The Economy as an Evolving Complex System II*. New York: Addison Wesley Longman.
- Ashley, R. (1989). 'Imposing International Purpose: Notes on a Problematique of Governance.' In *Global Changes and Theoretical Challenges: Approaches to World Politics for the 1990s*, eds E. Czempiel and J. Rosenau. Lexington, MA: Lexington.
- Austin, J. (1962). *How To Do Things With Words*. Oxford: Clarendon.
- Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, R. (1997). *The Complexity of Cooperation: Agent Based Models of Competition and Collaboration*. Princeton: Princeton University Press.

- Axtell, R. and Epstein, J. (1994). 'Agent based Modeling: Understanding Our Creations.' *Bulletin of the Santa Fe Institute* 9: 28–32.
- Bach, K. and R. Harnish (1979). *Linguistic Communication and Speech Acts*. Cambridge: MIT Press.
- Barry, C. A. (1998). 'Choosing Qualitative Data Analysis Software: Atlas/ti and Nudist Compared.' *Sociological Research Online* 3, <http://www.socresonline.org.uk/socresonline/3/3/4.html> (accessed 25 October 2006).
- Barthes, R. (1981). 'The Discourse of History.' *Comparative Criticism* 3: 7–20.
- Bates, R., A. Greif, M. Levi, J.-L. Rosenthal, and B. Weingast (1998). *Analytic Narratives*. Princeton: Princeton University Press.
- Bauman, Z. (1992). *Postmodern Ethics*. Oxford: Blackwell.
- Bell, V. and J. Butler (1999). 'On Speech, Race and Melancholia: An Interview with Judith Butler.' *Theory, Culture and Society* 16: 163–74.
- Bennett, A. and C. Elman (2007). 'Case Study Methods in the International Relations Subfield.' *Comparative Political Studies* 40: 170–95.
- Bererman, G. (1974). 'Bringing It All Back Home: Malaise in Anthropology.' In *Reinventing Anthropology*, ed. D. Hymes. New York: Vintage.
- Bererman, G. (1980). 'Are Human Rights Merely a Politicized Luxury in the World Today?' *Anthropology and Humanism Quarterly* 5: 2–13.
- Beyers, J. (2005). 'Multiple Embeddedness and Socialization in Europe: The Case of Council Officials.' *International Organization* 59: 899–936.
- Bigo, D. (2005). 'La Mondialisation de l' (in)sécurité? Réflexion sur le champ des professionnels de la gestion des inquiétudes et analytique de la transnationalisation des processus de l' (in)sécuritisation'. *Cultures & Conflits* 58: 53–100.
- Blakemore, D. (1992). *Understanding Utterances*. Oxford: Blackwell.
- Borofsky, R. (2005). *Yanomami: The Fierce Controversy and What We Can Learn From It*. Berkeley: University of California Press.
- Bourdieu, P. (1984). *Distinction: A Social Critique of Judgement of Taste*. Cambridge: Harvard University Press.
- Bourdieu, P. (1985). *In Other Words: Essays Toward a Reflexive Sociology*. Cambridge: Polity.
- Bourdieu, P. (1990). *The Logic of Practice*. Cambridge: Polity.
- Bourdieu, P. (1992). *Language and Symbolic Power*. Cambridge: Polity.
- Bourdieu, P. (1993). *The Field of Cultural Production: Essays on Art and Literature*. Cambridge: Polity.
- Bourdieu, P. (1995). *Outline of a Theory of Practice*. Cambridge: Cambridge University Press.
- Bourdieu, P. (1996a). *The Political Ontology of Martin Heidegger*. Cambridge: Polity.
- Bourdieu, P. (1996b). *The Rules of Art: Genesis and Structure of the Literary Field*. Cambridge: Polity.
- Bourdieu, P. (1998). *State Nobility: Elite Schools in the Field of Power*. Cambridge: Polity.
- Bourdieu, P. (1999). *The Weight of the World: Social Suffering in Contemporary Society*. Cambridge: Polity.
- Bourdieu, P. (2000a). *Pascalian Meditations*. Cambridge: Polity.
- Bourdieu, P. (2000b). *Homo Academicus*. Stanford: Stanford University Press.
- Bourdieu, P. (2002). *Practical Reason: On the Theory of Action*. Stanford: Stanford University Press.
- Bourdieu, P. (2004). *Science of Science and Reflexivity*. Chicago: University of Chicago Press.

- Bourdieu, P. (2005). *The Social Structures of the Economy*. Cambridge: Polity.
- Bourgois, P. (1995). *In Search of Respect: Selling Crack in El Barrio*. New York: Cambridge University Press.
- Brady, H. and D. Collier, eds (2004). *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham: Rowman and Littlefield.
- Brody, R., D. Mutz and P. Sniderman, eds (1996). *Political Persuasion and Attitude Change*. Ann Arbor: University of Michigan Press.
- Brown, C. (2005). *Postmodernism for Historians*. London: Pearson Longman.
- Brown, D. G., R. Riolo, D. T. Robinson, M. North, and W. Rand (2005). 'Spatial Process and Data Models: Toward Integration of Agent-Based Models and GIS'. *Journal of Geographical Systems* 7(1): 25–47.
- Brown, W. (1995) *States of Injury: Power and Freedom in Late Modernity*. Princeton: Princeton University Press.
- Campbell, D. (1993). *Politics Without Principles: Sovereignty, Ethics, and the Narratives of the Gulf War*. Boulder: Lynne Rienner.
- Caporaso, J., J. Checkel, and J. Jupille, eds (2003a). 'Integrating Institutions: Rationalism, Constructivism and the Study of the European Union.' Special Issue of *Comparative Political Studies* 36: 5–231.
- Caporaso, J., J. Checkel, and J. Jupille, eds (2003b). 'Introduction – Integrating Institutions: Rationalism, Constructivism and the Study of the European Union.' *Comparative Political Studies* 36: 7–41.
- Casti, J. (1997). 'Can You Trust It? On the Reliability of Computer Simulation and the Validity of Models.' *Complexity* 2: 8–11.
- Cederman, L. (1997). *Emergent Actors in World Politics*. Princeton: Princeton University Press.
- Cederman, L. (2001). 'Modeling the Democratic Peace as a Kantian Selection Process.' *Journal of Conflict Resolution* 45: 470–502.
- Cederman, L. (2003). 'Modeling the Size of Wars: From Billiard Balls to Sandpiles.' *American Political Science Review* 97: 135–50.
- Checkel, J. (2001). 'Why Comply? Social Learning and European Identity Change.' *International Organization* 55: 553–88.
- Checkel, J. (2003). 'Going Native' in Europe? Theorizing Social Interaction in European Institutions. *Comparative Political Studies* 36: 209–31.
- Checkel, J. (2005a). 'International Institutions and Socialization in Europe: Introduction and Framework.' *International Organization* 59: 801–26.
- Checkel, J. ed. (2005b). 'International Institutions and Socialization in Europe.' *International Organization* 59 (Special Issue).
- Checkel, J. (2006). 'Constructivism and EU Politics.' In *Handbook of European Union Politics*, eds K. E. Jørgensen, M. Pollack, and B. Rosamond. London: Sage.
- Chernoff, F. (2002). 'Scientific Realism as a Meta-Theory of International Politics.' *International Studies Quarterly* 46: 189–207.
- Chernoff, F. (2005). *The Power of International Theory: Reforging the Link to Foreign Policy-Making through Scientific Enquiry*. New York: Routledge.
- Clark, C., E. Sprenger, L. Vaneklasen, and L. Alpízar Durán (2006). *Where Is the Money for Women's Rights?* Washington, DC: Just Associates.
- Clifford, J. (1997). *Routes: Travel and Translation in the Late Twentieth Century*. Cambridge: Harvard University Press.
- Clifford, J. and G. E. Marcus, eds (1986). *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley: University of California Press.

- Cochran, M. (2002). 'Deweyan Pragmatism and Post-Positivist Social Science in IR.' *Millennium* 31: 525–48.
- Collier, D. and J. E. Mahon, Jr (1993). 'Conceptual "Stretching" Revisited: Adapting Categories in Comparative Analysis.' *American Political Science Review* 87: 845–55.
- Collins, P. H. (1991). *Black Feminist Thought: Knowledge, Consciousness, and the Politics of Empowerment*. Boston: Unwin Hyman.
- Comaroff, J. and J. Comaroff (1991). *Of Revelation and Revolution: Christianity, Colonialism, and Consciousness in South Africa*. Chicago: University of Chicago Press.
- Congressional Record, 12 August 1960: 16281; 17 August 1960: 16641; and 16 February 1961: A955.
- Conte, R., R. Hegselmann, and P. Terna, eds (1997). *Simulating Social Phenomena*. New York: Springer-Verlag.
- Council of Europe (1997). *European Convention on Nationality and Explanatory Report*. Strasbourg: Document DIR/JUR (97) 6 (14 May).
- Council of Europe (2000). *1st European Conference on Nationality: Trends and Developments in National and International Law on Nationality*. Strasbourg, 18–19 October 1999. Strasbourg: CONF/NAT (99) PRO 1 (3 February).
- Cox, R. (1981). 'Social Forces, States, and World Orders: Beyond International Relations Theory.' *Millennium* 10: 126–55.
- Crawford, N. (2002). *Argument and Change in World Politics: Ethics, Decolonization, and Humanitarian Intervention*. Cambridge: Cambridge University Press.
- D'Costa, B. (2006). 'Marginalized Identity: New Frontiers of Research for IR?' In *Feminist Methodologies for International Relations*, eds B. Ackerly, M. Stern, and J. True. Cambridge: Cambridge University Press.
- Derlugian, G. (2005). *Bourdieu's Secret Admirer in the Caucasus: A World System Biography*. Chicago: Chicago University Press.
- Dever, C. (2004). *Skeptical Feminism: Activist Theory, Activist Practice*. Minneapolis: University of Minnesota Press.
- Doty, R. (1996). *Imperial Encounters*. Minneapolis: University of Minnesota Press.
- Drezner, D. (2006). 'Notes from a Generalist.' *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 4 (Spring): 34–7.
- Driver, M. (1977). 'Individual Differences as Determinants of Aggression in the Inter-Nation Simulation.' In *A Psychological Examination of Political Leaders*, ed. M. G. Hermann. New York: Free Press.
- Duffy, G. and E. Goh (2007). 'Testing Sincerity: Kissinger's Opening Encounter with the Chinese Leadership.' *Journal of Language and Politics*. In press.
- Duffy, G., B. Frederking, and S. Tucker (1998). 'Language Games: Dialogical Analysis of INF Negotiations.' *International Studies Quarterly* 42: 271–94.
- Dunn, K. (2003). *Imagining the Congo*. New York: Palgrave Macmillan.
- Economist* (2002). 'The Brussels Consensus.' 7 December.
- Economist* (2003). 'Cracks in the College.' 13 September.
- van Eemeren, F. and R. Grootendorst (2004). *A Systematic Theory of Argumentation: The Pragma-dialectical Approach*. Cambridge: Cambridge University Press.
- van Eemeren, F., R. Grootendorst, S. Jackson, and S. Jacobs (1993). *Reconstructing Argumentative Discourse*. Tuscaloosa: University of Alabama Press.
- Ehrenreich, B. (2002). *Nickel and Dimed: On (Not) Getting By in America*. New York: Owl Books.

- Elman, C. and M. F. Elman (2001). *Bridges and Boundaries: Historians, Political Scientists, and the Study of International Relations*. Cambridge: MIT Press.
- Elster, J. (1998). 'A Plea for Mechanisms.' In *Social Mechanisms: An Analytical Approach to Social Theory*, eds P. Hedstroem and R. Swedberg. Cambridge: Cambridge University Press.
- Enloe, C. (1993). *The Morning After: Sexual Politics at the End of the Cold War*. Berkeley: University of California Press.
- Enloe, C. (2000). *Maneuvers: The International Politics of Militarizing Women's Lives*. Berkeley: University of California Press.
- Enloe, C. (2004). *The Curious Feminist: Searching for Women in a New Age of Empire*. Berkeley: University of California Press.
- Epstein, J. (1999). 'Agent-based Models and Generative Social Science.' *Complexity* 4: 1-60.
- Epstein, J. and R. Axtell (1996). *Growing Artificial Societies: Social Science from the Ground Up*. Washington, DC: Brookings Institution.
- Erikson, E. (1963). *Childhood and Society*. New York: Norton.
- Fearon, J. (1991). 'Counterfactuals and Hypothesis Testing in Political Science.' *World Politics* 43: 169-95.
- Fearon, J. and A. Wendt (2002). 'Rationalism v. Constructivism: A Skeptical View.' In *Handbook of International Relations*, eds W. Carlsnaes, T. Risse, and B. Simmons. London: Sage.
- Ferneu, E. (1969). *Guests of the Sheikh: An Ethnography of an Iraqi Village*. New York: Anchor Books.
- Finnemore, M. (1996). *National Interests in International Society*. Ithaca: Cornell University Press.
- Finnemore, M. and K. Sikkink (1998). 'International Norm Dynamics and Political Change.' *International Organization* 52: 887-918.
- Fluehr-Lobban, C. (1998). 'Ethics.' In *Handbook of Methods in Cultural Anthropology*, ed. H. Bernard. Walnut Creek, CA: Altamira Press.
- Fluehr-Lobban, C. (2003). *Ethics and the Profession of Anthropology: A Dialogue for Ethically Conscious Practice*. Walnut Creek, CA: Altamira Press.
- Forum (2005). 'Responses to Rosato, "The Flawed Logic of the Democratic Peace Theory."' *American Political Science Review* 99: 453-72.
- Fraser, A. (1999). 'Becoming Human: The Origins and Development of Women's Human Rights.' *Human Rights Quarterly* 21: 853-906.
- Fraser, N. (1997). *Justice Interruptus: Critical Reflections on The 'Postsocialist' Condition*. New York: Routledge.
- Frederking, B. (2000). *Resolving Security Dilemmas: A Constructivist Explanation of the INF Treaty*. Aldershot, UK: Ashgate.
- Fujimura, J. (1996). *Crafting Science: A Sociohistory of the Quest for the Genetics of Cancer*. Cambridge: Harvard University Press.
- Gaskins, R. (1992). *Burdens of Proof in Modern Discourse*. New Haven: Yale University Press.
- Gazdar, G. (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. New York: Academic Press.
- Geertz, C. (1968). *Islam Observed: Religious Development in Morocco and Indonesia*. New Haven: Yale University Press.
- Geertz, C. (1973). *The Interpretation of Cultures*. New York: Basic Books.

- George, A. (1979). 'Case Studies and Theory Development: The Method of Structured, Focused Comparison.' In *Diplomacy: New Approaches in History, Theory, and Policy*, ed. Paul Gordon Lauren. New York: Free Press.
- George, A. and A. Bennett (2005). *Case Studies and Theory Development in the Social Sciences*. Cambridge: MIT Press.
- Gerring, J. (2007). *Case Study Research: Principles and Practices*. New York: Cambridge University Press.
- Gheciu, A. (2005a). 'Security Institutions as Agents of Socialization? NATO and Post-Cold War Central and Eastern Europe.' *International Organization* 59: 973-1012.
- Gheciu, A. (2005b). *NATO in the 'New Europe': The Politics of International Socialization After the Cold War*. Stanford: Stanford University Press.
- Ginsburg, F. (1989). *Contested Lives: The Abortion Debate in an American Community*. Berkeley: University of California Press.
- Gluck, S. B. and D. Patai, eds (1991). *Women's Words: The Feminist Practice of Oral History*. New York: Routledge.
- Goertz, G. (2006). *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.
- Goodenough, W. (1981). *Culture, Language and Society*. Reading, MA: Addison-Wesley.
- Grice, H. P. (1957). 'Meaning.' *Philosophical Review*. 67. Reprinted in *Philosophical Logic*, ed. P. F. Strawson. Oxford: Oxford University Press, 1971.
- Gusterson, H. (1995a). 'Becoming a Weapons Scientist.' In *Technoscientific Imaginaries: Conversations, Profiles, and Memoirs*, ed. G. Marcus. Chicago: University of Chicago Press.
- Gusterson, H. (1995b). 'Short Circuit: Watching Television with a Nuclear Weapons Scientist.' In *The Cyborg Handbook*, ed. C. Gray. New York: Routledge.
- Gusterson, H. (1996). *Nuclear Rites: A Weapons Laboratory at the End of the Cold War*. Berkeley: University of California Press.
- Gusterson, H. (2003). 'Anthropology and the Military - 1968, 2003, and Beyond?' *Anthropology Today* 19 (3): 25-6.
- Gusterson, H. (2004). *People of the Bomb: Portraits of America's Nuclear Complex*. Minneapolis: University of Minnesota Press.
- Gusterson, H. (2005). 'The Ethics of Spying.' *Anthropology Today* 21 (4): 25.
- Guzzini, S. (2000). 'A Reconstruction of Constructivism in International Relations.' *European Journal of International Relations* 6: 147-82.
- Haas, P., ed. (1992). 'Knowledge, Power and International Policy Coordination.' *International Organization* 46 (Special Issue).
- Hacking, I. (1999). *The Social Construction of What?* Cambridge: Harvard University Press.
- Hall, J. R. (1999). *Cultures of Inquiry: From Epistemology to Discourse in Sociohistorical Research*. Cambridge: Cambridge University Press.
- Hall, P. (2003). 'Aligning Ontology and Methodology in Comparative Research.' In *Comparative Historical Analysis in the Social Sciences*, eds J. Mahoney and D. Rueschmeyer. New York: Cambridge University Press.
- Hamberg, S. (2005). 'From Causal Effects to Causal Mechanisms: Improving the Democratic Peace Propositions.' Unpublished manuscript, Department of Political Science, University of Oslo.

- Hansen, L. (1997). *Western Villains or Balkan Barbarism? Representations and Responsibility in the Debate over Bosnia*. PhD dissertation, Institut for Statskundskab, University of Copenhagen.
- Haraway, D. (1988). 'Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective.' *Feminist Studies* 14: 575-99.
- Harrell, M. (2003). 'Gender and Class-Based Role Expectations for Army Spouses.' In *Anthropology and the United States Military: Coming of Age in the Twenty-First Century*, eds P. R. Frese and M. C. Harrell. New York: Palgrave Macmillan.
- Harrison, N., ed. (2006). *Complexity in World Politics: Concepts and Methods of a New Paradigm*. Albany: SUNY Press.
- Hedstroem, P. and R. Swedberg, eds (1998). *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
- Hermann, M. (1977). 'Verbal Behavior of Negotiators in Periods of High and Low Stress: The 1965-66 New York City Transit Negotiations.' In *A Psychological Examination of Political Leaders*, ed. M. Hermann. New York: Free Press.
- Hermann, M. (1979). 'Indicators of Stress in Policy Makers during Foreign Policy Crises.' *Political Psychology* 1: 27-46.
- Hermann, M. (1980a). 'Explaining Foreign Policy Behavior Using Personal Characteristics of Political Leaders.' *International Studies Quarterly* 24: 7-46.
- Hermann, M. (1980b). 'Assessing the Personalities of Members of the Soviet Politburo.' *Personality and Social Psychology Bulletin* 6: 332-52.
- Hermann, M. (1984). 'Personality and Foreign Policy Decision Making: A Study of 53 Heads of Government.' In *Foreign Policy Decision Making: Perceptions, Cognition, and Artificial Intelligence*, eds D. Sylvan and S. Chan. New York: Praeger.
- Hermann, M. (1985). 'Validating a Technique for Assessing Personalities of Political Leaders at a Distance: A Test Focusing on Heads of Government.' Report prepared for Defense Systems, Inc. (Washington, DC).
- Hermann, M. (1986). 'Effects of Speech and Interview Materials on Profiles of Leaders at a Distance: A Validation Exercise.' Report prepared for Defense Systems, Inc. (Washington, DC).
- Hermann, M. (1988). 'Validating a Technique for Assessing Personalities of Political Leaders at a Distance: Profiles of 12 Leaders from the Same Culture.' Report prepared for Defense Systems, Inc. (Washington, DC).
- Hermann, M. (1999). 'Leadership Profile of Slobodan Milosevic.' *Social Science Automation*. Columbus: Ohio.
- Hermann, M. (2000). 'Leadership Profile of Jiang Zemin.' *Social Science Automation*. Columbus: Ohio.
- Hermann, M. (2005). 'William Jefferson Clinton's Leadership Style.' In *The Psychological Assessment of Political Leaders*, ed. J. Post. Ann Arbor: University of Michigan Press.
- Hermann, M. and C. Hermann (1989). 'Who Makes Foreign Policy and How: An Empirical Inquiry.' *International Studies Quarterly* 33: 361-87.
- Hermann, M. and C. Hermann (1990). 'Hostage-Taking, Stress, and the Presidency.' In *The Origins of Terrorism: Psychologies, Ideologies, Theologies, and States of Mind*, ed. W. Reich. Cambridge: Cambridge University Press.
- Hermann, M. and T. Preston (1998). 'Presidents, Leadership Style, and the Advisory Process.' In *Domestic Sources of American Foreign Policy*, eds J. McCormick and E. Wittkopf. New York: Rowman and Littlefield.
- Hirschmann, N. (2003). *The Subject of Liberty: Toward a Feminist Theory of Freedom*. Princeton: Princeton University Press.

- Hoffmann, M. (2003). 'Self-Organized Criticality and Norm Avalanches.' Paper presented at the annual meeting of the North American Association of Computational Social and Organizational Science, Pittsburgh, PA, June.
- Hoffmann, M. (2005a). *Ozone Depletion and Climate Change: Constructing a Global Response*. Albany: SUNY Press.
- Hoffmann, M. (2005b). 'Self-Organized Criticality and Norm Avalanches.' *Proceedings of the Symposium on Normative Multi-Agent Systems AISB*: 117–25.
- Hoffmann, M., H. Kelley, and T. Evans (2002). 'Simulating Land-Cover Change in South-Central Indiana: An Agent-Based Model of Deforestation and Afforestation.' In *Complexity and Ecosystem Management*, ed. M. Janssen. Cheltenham, UK: Edward Elgar.
- Holsti, O. R. (1969). *Content Analysis for the Social Sciences and the Humanities*. Reading, MA: Addison Wesley.
- Holsti, O. R. (1972). *Crisis, Escalation, War*. Montreal: McGill-Queens University Press.
- Hooghe, L. (2005). 'Several Roads Lead To International Norms, But Few via International Socialization: A Case Study of the European Commission.' *International Organization* 59: 861–98.
- Hopf, T. (2002). *Social Construction of International Politics: Identities and Foreign Policies, Moscow, 1955 and 1999*. Ithaca: Cornell University Press.
- Hopf, T. (2007). 'The Limits of Interpreting Evidence.' In *Theory and Evidence in Comparative Politics and International Relations*, eds R. N. Lebow and M. Lichbach. New York: Palgrave Macmillan.
- Hovi, J. (2004). 'Causal Mechanisms and the Study of International Environmental Regimes.' In *Regime Consequences: Methodological Challenges and Research Strategies*, eds A. Underdal and O. Young. Boston: Kluwer Academic Publishers.
- Hug, S. and T. Koenig (2000). 'Ratifying Maastricht: Parliamentary Votes on International Treaties and Theoretical Solution Concepts.' *European Union Politics* 1: 93–124.
- Hug, S. and T. Koenig (2002). 'In View of Ratification: Governmental Preferences and Domestic Constraints at the Amsterdam Intergovernmental Conference.' *International Organization* 56: 447–76.
- Hurwitz, R. (1989). 'Strategic and Social Fictions in the Prisoner's Dilemma.' In *International/Intertextual Relations*, eds J. Der Derian and M. Shapiro. Lexington, MA: Lexington Books.
- Hurwitz, R., J. C. Mallery, and G. Duffy (1987). 'Hermeneutics.' In *The Encyclopedia of Artificial Intelligence*, ed. S. Shapiro. New York: John Wiley and Sons.
- Hutchinson, S. (1996). *Nuer Dilemmas: Coping with Money, War and the State*. Berkeley: University of California Press.
- Jacoby, T. (2006). 'From the Trenches: Dilemmas of Feminist IR Fieldwork.' In *Feminist Methodologies for International Relations*, eds B. Ackerly, M. Stern, and J. True. Cambridge: Cambridge University Press.
- Janssen, M., ed. (2002). *Complexity and Ecosystem Management*. Cheltenham, UK: Edward Elgar.
- Johnson, J. (2002). 'How Conceptual Problems Migrate: Rational Choice, Interpretation and the Hazards of Pluralism.' *Annual Review of Political Science* 5: 223–48.
- Johnson, P. (1999). 'Simulation Modeling in Political Science.' *American Behavioral Scientist* 42: 1509–30.

- Johnson, P. (2006). 'Consequences of Positivism: A Pragmatist Assessment.' *Comparative Political Studies* 39: 224–52.
- Johnston, A. (2001). 'Treating International Institutions as Social Environments.' *International Studies Quarterly* 45: 487–515.
- Johnston, A. (2005). 'Conclusions and Extensions – Toward Mid-Range Theorizing and Beyond Europe.' *International Organization* 59: 1013–44.
- Johnston, A. (2007). *Social States: China in International Institutions*. Princeton: Princeton University Press.
- Kaarbo, J. and M. Hermann (1998). 'Leadership Styles of Prime Ministers: How Individual Differences Affect the Foreign Policymaking Process.' *Leadership Quarterly* 9: 9243–63.
- Kane, S. (1998). *AIDS Alibis: Sex, Drugs and Crime in the Americas*. Philadelphia: Temple University Press.
- Katzenstein, P., ed. (1996). *The Culture of National Security: Norms and Identity in World Politics*. New York: Columbia University Press.
- Katzenstein, P. and R. Sil (2005). 'What is Analytic Eclecticism and Why Do We Need It? A Pragmatist Perspective on Problems and Mechanisms in the Study of World Politics.' Paper presented at the Annual Convention of the American Political Science Association (www.apsanet.org).
- Keller, J. (2005). 'Constraint Challengers, Constraint Respecters, and Foreign Policy.' *Political Psychology* 26: 835–67.
- Kelley, J. (2004). *Ethnic Politics in Europe: The Power of Norms and Incentives*. Princeton: Princeton University Press.
- Keohane, R. (2001). 'Governance in a Partially Globalized World.' *American Political Science Review* 95: 1–13.
- King, G., R. Keohane, and S. Verba (1994). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Klotz, A. (1995). *Norms in International Relations: The Struggle against Apartheid*. Ithaca: Cornell University Press.
- Klotz, A. (2002). 'Transnational Activism and Global Transformations: The Anti-Apartheid and Abolitionist Experiences.' *European Journal of International Relations* 8: 49–76.
- Klotz, A. and C. Lynch (2006). 'Translating Terminologies,' *International Studies Review* 8: 356–62.
- Klotz, A. and C. Lynch (2007). *Strategies for Research in Constructivist International Relations*. Armonk, NY: M. E. Sharpe.
- Klotz, I. (1986). *Diamond Dealers and Feather Merchants: Tales from the Sciences*. Boston: Birkhäuser.
- Knorr Cetina, K. (1999). *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge: Harvard University Press.
- Kollman, K., J. Miller, and S. Page (1997). 'Computational Political Economy.' In *The Economy as an Evolving Complex System II*, eds B. Arthur, S. Durlauf, and D. Lane. Reading, MA: Addison Wesley.
- Koselleck, R. (1988). *Critique and Crisis: Enlightenment and the Pathogenesis of Modern Society*. Cambridge: MIT Press.
- Kratochwil, F. and J. G. Ruggie (1986). 'International Organization: A State of the Art on An Art of the State.' *International Organization* 40: 753–75.
- Laclau, E. and C. Mouffe (1985). *Hegemony and Socialist Strategy: Towards a Radical Democratic Politics*. London: Verso.

- Laffan, B. (1998). 'The European Union: A Distinctive Model of Internationalization.' *Journal of European Public Policy* 5: 235–53.
- Lahire, B., ed. (1999). *Le travail sociologique de Pierre Bourdieu. Dettes et critiques*. Paris: La Découverte.
- Lane, R. (1996). 'Positivism, Scientific Realism and Political Science: Recent Developments in the Philosophy of Science.' *Journal of Theoretical Politics* 8(3): 361–82.
- Lasswell, H. (1936). *Politics: Who Gets What, When, How*. New York: McGraw-Hill.
- Latour, B. and S. Woolgar (1986). *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.
- Lazreq, M. (2002). 'Development: Feminist Theory's Cul-De-Sac.' In *Feminist Post-Development Thought: Rethinking Modernity, Postcolonialism & Representation*, ed. K. Saunders. London: Zed.
- Leander, A. (2001). 'Pierre Bourdieu on Economics.' *Review of International Political Economy* 8: 344–53.
- Leander, A. (2002). 'Do we really need reflexivity in IPE? Bourdieu's two reasons for answering affirmatively (contribution to a colloquium on Pierre Bourdieu).' *Review of International Political Economy* 9: 601–9.
- Lebow, R. N. and M. Lichbach, eds (2007). *Theory and Evidence in Comparative Politics and International Relations*. New York: Palgrave Macmillan.
- Lee, J. (1986). 'Steven Mosher vs. Stanford.' *The Monthly*, April: 33–40.
- Lefcourt, H. (1976). *Locus of Control: Current Trends in Theory and Research*. New York: Halstead.
- Lévi-Strauss, C. (1963). *Structural Anthropology*. New York: Basic Books.
- Levinson, D. (1978). *The Seasons of a Man's Life*. New York: Ballantine.
- Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, S. (2000). *Presumptive Meanings: The Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.
- Lewis, J. (2005). 'The Janus Face of Brussels: Socialization and Everyday Decision-Making in the European Union.' *International Organization* 59: 937–71.
- Ling, L. H. M. (2002). *Postcolonial International Relations: Conquest and Desire between Asia and the West*. New York: Palgrave.
- Little, D. (1991). *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder: Westview.
- Luhrmann, T. (2001). *Of Two Minds: An Anthropologist Looks at American Psychiatry*. New York: Vintage.
- Lukes, S. (1974). *Power: A Radical View*. London: Macmillan.
- Lustick, I. (2000). 'Agent-based Modeling of Collective Identity: Testing Constructivist Theory.' *Journal of Artificial Societies and Social Simulation* 3, www.soc.surrey.ac.uk/JASSS/3/1/1.html.
- Lustick, I., D. Miodownik, and R. Eidelson (2004). 'Secessionism in Multicultural States: Does Sharing Power Prevent or Encourage It?' *American Political Science Review* 98: 209–29.
- MacDonald, P. (2003). 'Useful Fiction or Miracle Maker: The Competing Epistemological Foundations of Rational Choice Theory.' *American Political Science Review* 97: 551–65.
- MacKinnon, C. (2006). *Are Women Human? And Other International Dialogues*. Cambridge: Belknap Press of Harvard University Press.
- Macy, M. and R. Willer (2002). 'From Factors to Actors: Agent-based Modeling and Computational Sociology.' *Annual Review of Sociology* 48: 143–67.

- Magnetie, P. (2004). 'Coping with Constitutional Incompatibilities: Bargains and Rhetoric in the Convention on the Future of Europe.' Paper presented at the ARENA Research Seminar. Oslo: ARENA Centre for European Studies, University of Oslo.
- Mahoney, J. and G. Goertz (2004). 'The Possibility Principle: Choosing Negative Cases in Comparative Research,' *American Political Science Review* 98: 653–99.
- Majeski, S. (2004). 'Asymmetric Power Among Agents and the Generation and Maintenance of Cooperation in International Relations.' *International Studies Quarterly* 48: 455–70.
- Majeski, S. (2005). 'Do Exploitive Agents Benefit from Asymmetric Power to International Politics?' *British Journal of Political Science* 35: 745–55.
- Mallery, J. C. and F. L. Sherman (1993). 'Learning Historical Rules of Major Power Intervention in the Post-War International System.' Paper presented to the International Studies Association.
- Marcus, G. (1995). 'Ethnography in/of the World System: The Emergence of Multi-Sited Ethnography.' *Annual Review of Anthropology* 24: 95–117.
- Marcus, G. and M. Fischer (1986). *Anthropology as Cultural Critique*. Chicago: University of Chicago Press.
- Mayntz, R. (2003). 'Mechanisms in the Analysis of Macro-Social Phenomena.' MPIfG Working Paper, 03/3. Cologne: Max Planck Institute for the Study of Societies (April).
- McClelland, D. (1975). *Power: The Inner Experience*. New York: Irvington.
- McFate, M. (2005). 'The Military Utility of Understanding Adversary Culture.' *Joint Force Quarterly* 38: 42–8.
- Mey, J. (2001). *Pragmatics: An Introduction*, 2nd edn. London: Blackwell.
- Milliken, J. (1999). 'The Study of Discourse in International Relations: A Critique of Research and Methods.' *European Journal of International Relations* 5: 225–54.
- Monroe, K. R., ed. (2005). *Perestroika: The Raucous Rebellion in Political Science*. New Haven: Yale University Press.
- Moos, F. (2005a). 'Anthropologists as Spies.' *Anthropology Today* 21 (3): 25–6.
- Moos, F. (2005b). 'Some Thoughts on Anthropological Ethics and Today's Conflicts.' *Anthropology Newsletter* 46 (6): 40–2.
- Moravcsik, A. (1999). 'Is Something Rotten in the State of Denmark? Constructivism and European Integration.' *Journal of European Public Policy* 5: 669–81.
- Moyser, G. and M. Wagstaffe (1987). *Research Methods for Elite Studies*. London: Allen and Unwin.
- Myerhoff, B. (1976). *Peyote Hunt: The Sacred Journey of the Huichol Indians*. Ithaca: Cornell University Press.
- Nader, L. (1974). 'Up the Anthropologist: Perspectives Gained from Studying Up.' In *Reinventing Anthropology*, ed. D. Hymes. New York: Vintage.
- Nahavandi, A. (2003). *The Art and Science of Leadership*, 3rd edn. Upper Saddle River, NJ: Prentice-Hall.
- Narayan, U. (1997). *Dislocating Cultures: Identities, Traditions, and Third-World Feminism*. New York: Routledge.
- National Science Foundation, Political Science Program, Directorate for Social, Behavioral, and Economic Sciences (2002). *Empirical Implications of Theoretical Models Report*. Washington, DC: National Science Foundation.
- Neumann, I. (1996). *Russia and the Idea of Europe. A Study in Identity and International Relations*. London: Routledge.

- Neumann, I. (1997). 'Book Review Essay: Ringmar on Identity and War.' *Cooperation and Conflict* 32: 309–30.
- Neumann, I. (1999). *Uses of the Other: 'The East' in European Identity Formation*. Minneapolis: University of Minnesota Press.
- Neumann, I. (2001). *Norge – en kritikk. Begrepsmakt i Europa-debatten*. Oslo: Pax.
- Neumann, I. (2002). 'This Little Piggy Stayed at Home: Why Norway is Not a Member of the EU.' In *European Integration and National Identity: The Challenge of the Nordic States*, eds L. Hansen and O. Wæver. London: Routledge.
- Neumann, I. (2004). 'Deep Structure, Free-Floating Signifier, or Something in Between? Europe's Alterity in Putin's Russia.' In *Identity and Global Politics: Empirical and Theoretical Elaborations*, eds P. M. Goff and K. C. Dunn. London: Palgrave.
- Neumann, I. (2007). 'A Speech that the Entire Ministry May Stand For.' *International Political Sociology* 1: 183–200.
- Neumann, I. and D. Nexon (2006). 'Harry Potter and the Study of World Politics.' In *Harry Potter and Global Politics*, eds D. Nexon and I. B. Neumann. Lanham, MD: Rowman and Littlefield.
- Newbury, D. (1998). 'Understanding Genocide.' *African Studies Review* 41: 73–97.
- Northstrom, C. (2004). *Shadows of War: Violence, Power, and International Profiterring in the Twenty-First Century*. Berkeley: University of California Press.
- Nussbaum, M. (2001). *Upheavals of Thought: The Intelligence of Emotions*. Cambridge: Cambridge University Press.
- Odell, J. (2006). 'A Major Milestone with One Major Limitation.' *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods* 4 (Spring): 37–40.
- Orbell, J., R. Dawes, and A. Van de Kragt (1988). 'Explaining Discussion-Induced Cooperation.' *Journal of Personality and Social Psychology* 34: 811–19.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz (1994). 'Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences.' *Science* 263: 641–46.
- Page, S. (1999). 'Computational Models from A to Z.' *Complexity* 5: 35–41.
- Parker, D. (2005). 'Integration of Geographic Information Systems and Agent-Based Models of Land Use: Challenges and Prospects.' In *GIS, Spatial Analysis and Modeling*, eds D. Maguire, M. Goodchild, and M. Batty. Redlands, CA: ESRI Press.
- Parker, D., S. Manson, M. Janssen, M. J. Hoffmann, and P. Deadman (2003). 'Multi-Agent Systems for the Simulation of Land-Use and Land-Cover Change: A Review.' *Annals of the Association of American Geographers* 93: 314–37.
- Pateman, C. (1988). *The Sexual Contract*. Stanford: Stanford University Press.
- Pepinsky, T. (2005). 'From Agents to Outcomes: Simulation in International Relations.' *European Journal of International Relations* 11: 367–94.
- Perloff, R. (1993). *The Dynamics of Persuasion*. Hillsdale, NJ: Erlbaum Associates.
- Peterson, S. and M. J. Tierney with D. Maliniak (2005). 'Teaching and Research Practices, Views on the Discipline, and Policy Attitudes of International Relations Faculty at US Colleges and Universities.' College of William and Mary, <http://www.wm.edu/irtheoryandpractice/trip/papers.php>.
- Peterson, V. (1990). 'Whose Rights? A Critique of the "Givens" in Human Rights Discourse.' *Alternatives* 15: 303–44.
- Pettman, J. (1992). *Living in the Margins: Racism, Sexism, and Feminism in Australia*. North Sydney, Australia: Allen and Unwin.

- Post, J. (1980). 'The Seasons of a Leader's Life: Influences on the Life cycle in Political Behaviour.' *Political Psychology* 2: 35-49.
- Post, J. (1993). 'Current Concepts of Narcissism: Implications for Political Psychology.' *Political Psychology* 14: 99-121.
- Post, J. (2005). *The Psychological Assessment of Political Leaders*. Ann Arbor: University of Michigan Press.
- Pouliot, V. (2003). 'La Russie et la communauté atlantique. Vers une culture commune de sécurité?' *Études Internationales* 34: 25-51.
- Pouliot, V. (2007). 'Subjectivism: Toward a Constructivist Methodology.' *International Studies Quarterly* 51: 359-84.
- Preston, T. (2001). *The President and His Inner Circle: Leadership Style and the Advisory Process in Foreign Policy*. New York: Columbia University Press.
- Preston, T. and M. Hermann (2003). 'Presidential Leadership Style and the Foreign Policy Advisory Process.' In *The Domestic Sources of American Foreign Policy*, eds E. Wittkopf and J. McCormick. New York: Rowman and Littlefield.
- Price, D. (2000). 'Anthropologists as Spies.' *The Nation*, 2 November.
- Price, D. (2004). *Threatening Anthropology: McCarthyism and the FBI's Surveillance of Activist Anthropologists*. Durham: Duke University Press.
- Price, D. and H. Gusterson (2005). 'Spies in Our Midst.' *Anthropology Newsletter* 46 (6): 39-40.
- Price, R. (1997). *The Chemical Weapons Taboo*. Ithaca: Cornell University Press.
- Ragin, C. (2000). *Fuzzy-set Social Science*. Chicago: University of Chicago Press.
- Reich, W. (1933). *Character Analysis*. New York: Orgone Institute Press.
- Rescher, N. (1977). *Dialectics*. Albany: SUNY Press.
- Ringmar, E. (1996). *Identity, Interest and Action: A Cultural Explanation of Sweden's Intervention in the Thirty Years War*. Cambridge: Cambridge University Press.
- Risman, B. (2004). 'Gender as a Social Structure: Theory Wrestling with Activism.' *Gender & Society* 18: 429-50.
- Rosaldo, R. (1989). *Culture and Truth: The Remaking of Social Analysis*. Boston: Beacon Press.
- Rosato, S. (2003). 'The Flawed Logic of Democratic Peace Theory.' *American Political Science Review* 97: 585-602.
- Rotter, J. (1966). 'Generalized Expectancies for Internal Versus External Control of Reinforcement.' *Psychological Monographs* 80 (Whole No. 609).
- Rousseau, D. and M. van der Veen (2005). 'The Emergence of a Shared Identity: An Agent-Based Computer Simulation of Idea Diffusion.' *Journal of Conflict Resolution* 49: 686-712.
- Ruggie, J. (1998). *Constructing the World Polity: Essays on International Institutionalization*. London: Routledge.
- Sahlins, M. (2000). 'Review of James Tierney's *Darkness in El Dorado*.' *Washington Post Book World*, December 10 (1).
- Said, E. (1978). *Orientalism*. Harmondsworth: Penguin.
- Sandoval, C. (2000). *Methodology of the Oppressed*. Minneapolis: University of Minnesota Press.
- Sanjek, R. (1990). *Fieldnotes: The Makings of Anthropology*. Ithaca: Cornell University Press.
- Sartori, G. (1970). 'Concept Misinformation in Comparative Politics.' *American Political Science Review* 64: 1033-53.
- Schafer, M. (2000). 'Issues in Assessing Psychological Characteristics at a Distance: An Introduction to the Symposium.' *Political Psychology* 21: 511-27.

- Schimmelfennig, F. (2003). *The EU, NATO and the Integration of Europe: Rules and Rhetoric*. Cambridge: Cambridge University Press.
- Schimmelfennig, F. (2005). 'Strategic Calculation and International Socialization: Membership Incentives, Party Constellations, and Sustained Compliance in Central and Eastern Europe.' *International Organization* 59: 827-60.
- Schwartz-Shea, P. (2005). 'The Graduate Student Experience: "Hegemony" or Balance in Methodological Training?' In *Perestroika: The Raucous Rebellion in Political Science*, ed. K. Monroe. New Haven: Yale University Press.
- Searle, J. (1969). *Speech Acts*. Cambridge: Cambridge University Press.
- Sherman, F. L. (1994). 'SherFACS: A Cross-Paradigm, Hierarchical and Contextually Sensitive Conflict Data Set.' *International Interactions* 20: 79-100.
- Sherman, F. L., J. C. Mallery, S. C. Unseld, and G. Duffy (1992). 'Grammars of Conflict and Cooperation in the Post-War International System.' Paper presented to the International Studies Association.
- Shusterman, R., ed. (1999). *Bourdieu: A Critical Reader*. Oxford: Blackwell.
- Simon, M. and H. Starr (2000). 'Two-Level Security Management and the Prospects for New Democracies: A Simulation Analysis.' *International Studies Quarterly* 44: 391-422.
- Smith, L. (1999). *Decolonizing Methodologies: Research and Indigenous Peoples*. London: Zed.
- Sprinz, D. and Y. Wolinsky-Nahmias (2004). "Introduction: Methodology in International Relations Research." In *Models, Numbers and Cases: Methods for Studying International Relations*, eds D. Sprinz and Y. Wolinsky-Nahmias. Ann Arbor: University of Michigan Press.
- Stacey, J. (1999). 'Ethnography Confronts the Global Village.' *Journal of Contemporary Ethnography* 28: 687-97.
- Stack, C. (1997). *All Our Kin*, 2nd edn. New York: Basic Books.
- Staeheli, L. and R. Nagar (2002). 'Feminists Talking across Worlds.' *Gender, Place and Culture* 9: 167-72.
- Steinmetz, G., ed. (2005). *The Politics of Method in the Human Sciences: Positivism and Its Epistemological Others*. Durham: Duke University Press.
- Stern, M. (2006). 'Racism, Sexism, Classism and Much More: Reading Security-Identity in Marginalized Sites.' In *Feminist Methodologies for International Relations*, eds B. Ackerly, M. Stern, and J. True. Cambridge: Cambridge University Press.
- Stolberg, S. (1999). 'The Biotech Death of Jesse Gelsinger.' *New York Times Magazine*, November 28.
- Stoller, P. (1989). *The Taste of Ethnographic Things: The Senses in Anthropology*. Philadelphia: University of Pennsylvania Press.
- Suedfeld, P. (1992). 'Cognitive Misers and Their Critics.' *Political Psychology* 13: 435-53.
- Suedfeld, P. and M. Wallace (1995). 'President Clinton as a Cognitive Manager.' In *The Clinton Presidency: Campaigning, Governing and the Psychology of Leadership*, ed. S. Renshon. Boulder: Westview Press.
- Sun, M. (1983). 'The Mysterious Expulsion of Steven Mosher.' *Science* 220 (May 13): 692.
- Suu Kyi, A. (1991). *Freedom from Fear: And Other Writings*. New York: Penguin.
- Sylvester, C. (1994). *Feminist Theory and International Relations in a Postmodern Era*. Cambridge: Cambridge University Press.

- Tannenwald, N. (2007). *The Nuclear Taboo: The United States and the Non-Use of Nuclear Weapons since 1945*. Cambridge: Cambridge University Press.
- Tetlock, P. (1991). *An Integratively Complex Look at Integrative Complexity*. Paper presented at the annual meeting of the American Psychological Association, San Francisco, August.
- Tetlock, P. and A. Belkin (1996). *Counterfactual Thought Experiments in World Politics*. Princeton: Princeton University Press.
- Tierney, P. (2000). *Darkness in El Dorado: How Scientists and Journalists Devastated the Amazon*. New York: Norton.
- Tilly, C. (1984). *Big Structures, Large Processes, Huge Comparisons*. New York: Russell Sage Foundation.
- Tilly, C. (2001). 'Mechanisms in Political Processes.' *Annual Review of Political Science* 4: 21–41.
- Torring, J. (1999). *New Theories of Discourse: Laclau, Mouffe and Zizek*. Oxford: Blackwell.
- Trachtenberg, M. (2006). *The Craft of International History: A Guide to Method*. Princeton: Princeton University Press.
- Turner, W. (1983). 'Stanford Ousts a PhD. Candidate over his use of data on China.' *New York Times*, February 26 (A7).
- Unsel, S. and J. Mallery (1991). 'Inductive Interaction Detection for Complex Datamodels.' MIT Artificial Intelligence Laboratory Memo 1298.
- Vaillant, G. (1992). *Ego Mechanisms of Defense: A Guide for Clinicians and Researchers*. Washington, DC. American Psychiatric Press.
- Vélez-Ibañez, C. (1983). *Rituals of Marginality: Politics, Process, and Culture Change in Urban Central Mexico, 1969-1974*. Berkeley: University of California Press.
- Wacquant, L. (2003). *Body and Soul: Notebooks of an Apprentice Boxer*. New York: Oxford University Press.
- Wacquant, L., ed. (2005). *Pierre Bourdieu and Democratic Politics: The Mystery of Ministry*. Cambridge: Polity.
- Wæver, O. (1999). 'Identity, Community and Foreign Policy: Discourse Analysis as Foreign Policy Theory.' In *Between Nations and Europe: Regionalism, Nationalism and the Politics of Union*, eds L. Hansen and O. Wæver. London: Routledge.
- Wakin, E. (1992). *Anthropology Goes to War: Professional Ethics and Counterinsurgency in Thailand*. Madison: University of Wisconsin Center for Southeast Asian Studies Monograph.
- Walker, S. (1983). 'The Motivational Foundations of Political Belief Systems: A Re-Analysis of the Operational Code Construct.' *International Studies Quarterly* 27: 179–201.
- Wax, M. (2003). 'Wartime Dilemmas of an Ethical Anthropology.' *Anthropology Today* 19 (3): 23–4.
- Weber, C. (1995). *Simulating Sovereignty: Intervention, the State, and Symbolic Exchange*. Cambridge: Cambridge University Press.
- Weber, C. (1999). *Faking It: U.S. Hegemony in a 'Post-Phallic' Era*. Minneapolis: University of Minnesota Press.
- Weber, R. (1990). *Basic Content Analysis*. Newbury Park: Sage.
- Weber, S. (1994). 'Origins of the European Bank for Reconstruction and Development.' *International Organization* 48: 1–38.
- Weldes, J. (1999). *Constructing National Interests: The United States and the Cuban Missile Crisis*. Minneapolis: University of Minnesota Press.

- Weldes, J., M. Laffey, H. Gusterson, and R. Duvall (1999). 'Introduction: Constructing Insecurity.' In *Cultures of Insecurity. States, Communities, and the Production of Danger*, eds J. Weldes, M. Laffey, H. Gusterson, and R. Duvall. Minneapolis: University of Minnesota Press.
- Wendt, A. (1999). *Social Theory of International Politics*. Cambridge: Cambridge University Press.
- Wessels, W. (1998). 'Comitology: Fusion in Action – Politico-Administrative Trends in the EU System.' *Journal of European Public Policy* 5: 209–34.
- White, H. (1978). *Tropics of Discourse: Essays in Cultural Criticism*. Baltimore: John Hopkins University Press.
- Wight, C. (2002). 'Philosophy of Social Science and International Relations.' In *Handbook of International Relations*, eds W. Carlsnaes, T. Risse, and B. Simmons. London: Sage.
- Wilson, J. (1990). *Politically Speaking: The Pragmatic Analysis of Political Language*. Oxford: Basil Blackwell.
- Winter, D. (1973). *The Power Motive*. New York: Free Press.
- Winter, D. (1992). 'Personality and Foreign Policy: Historical Overview.' In *Political Psychology and Foreign Policy*, eds E. Singer and V. Hudson. Boulder: Westview.
- Winter, D. (1995). 'Presidential Psychology and Governing Styles: A Comparative Psychological Analysis of the 1992 Presidential Candidates.' In *The Clinton Presidency: Campaigning, Governing and the Psychology of Leadership*, ed. S. Renshon. Boulder: Westview.
- Winter, D. (2005). 'Measuring the Motives of Political Actors at a Distance.' In *The Psychological Assessment of Political Leaders*, ed. J. Post. Ann Arbor: University of Michigan Press.
- Winter, D. and A. Stewart (1977). 'Content Analysis as a Technique for Assessing Political Leaders.' In *A Psychological Examination of Political Leaders*, ed. M. Hermann. New York: Free Press.
- Winter, D., M. Hermann, W. Weintraub, and S. G. Walker (1991). 'The personalities of Bush and Gorbachev Measured at a Distance: Procedures, Portraits, and Policy.' *Political Psychology* 12: 215–43.
- Wolf, D., ed. (1996). *Feminist Dilemmas in Fieldwork*. Boulder: Westview.
- Wolf, M. (1992). *A Thrice-Told Tale: Feminism, Postmodernism, and Ethnographic Responsibility*. Stanford: Stanford University Press.
- World Social Forum Panel (2004a). 'A Debate on International Issues among Social Movements.' In *World Social Forum, Centre for Equity Studies, January 18*. Mumbai, India.
- World Social Forum Panel (2004b). 'A Dialogue between Movements: Breaking Barriers and Building Bridges.' In *World Social Forum, National Network of Autonomous Women's Groups (DAWN, AFM, and WICEJ), January 19*. Mumbai, India.
- World Social Forum Panel (2004c). 'A Dialogue between Various Movements on Sexuality Issues.' In *World Social Forum, Rainbow Planet, January 17*. Mumbai, India.
- Yanow, D. and P. Schwartz-Shea, eds (2006). *Interpretation and Method: Empirical Research Methods and the Interpretive Turn*. Armonk, NY: M. E. Sharpe.
- Young, I. (1990). *Justice and the Politics of Difference*. Princeton: Princeton University Press.

- Young, I. (2001). 'Activist Challenges to Deliberative Democracy.' *Political Theory* 29: 670–90.
- Young, M. (2000). 'Automating Assessment at a Distance.' *The Political Psychologist* 5: 17–23.
- Young, M. and M. Schafer (1998). 'Is There Method in Our Madness? Ways of Assessing Cognition in International Relations.' *Mershon International Studies Review* 42: 63–96.
- Zehfuss, M. (2002). *Constructivism in International Relations: The Politics of Reality*. Cambridge: Cambridge University Press.
- Zimbardo, P. and M. Leippe (1991). *The Psychology of Attitude Change and Social Influence*. New York: McGraw Hill.
- Zürn, M. and J. Checkel (2005). 'Getting Socialized to Build Bridges: Constructivism and Rational Choice, Europe and the Nation State.' *International Organization* 59: 1045–79.

Index

- Ackerly (Brooke), 24, 46, 48, 51, 65,
69, 82, 87, 90, 98, 225, 227
see also feminism, reflexivity
- actors/action, *see* agency (agents)
- agency (agents), 5–6, 17–19, 21–2, 62,
75, 76, 78, 80–5, 114, 115–16,
117, 124, 168, 199–200
see also agent-based modeling,
personality, rationality
- agent-based modeling (ABM), 3, 6–7,
123, 187–208, 220, 223, 226, 227
see also modeling
- anthropology, 5, 93–5, 222
see also ethnography, social sciences
- archaeology, 74
see also genealogy
- archives, *see* historiography
- assignments, *see* courses
- Barkin (Samuel), 222, 228, 231, 233,
234–5, 236
see also courses,
qualitative-quantitative
distinction, pedagogy
- bias, *see* cultural competence, validity
- biography, *see* data, narrative
- Bourdieu (Pierre), 4, 11–27, 70–1, 82,
226
see also conceptualization,
Leander(Anna)
- cartoons, *see* culture, data, texts
- case studies, 4, 43–58, 85–6, 116, 147,
202, 213–15
non-cases, 46–7
small-n, 54–7, 213, 215
universe of cases, 43, 46
see also comparison
- categorization, 25, 27, 211–20
see also conceptualization
- causality, 2, 3, 5, 48, 49–50, 56, 58, 78,
84–5, 114, 115–16, 123, 175, 184
see also constitution, epistemology,
interpretation
- censorship, 5, 66, 70, 72, 84
- Checkel (Jeffrey), 49, 85, 105, 192,
194, 196, 199–200, 213, 214, 236
see also causality, process tracing
- coding, 2, 6, 151, 157–60, 179, 185–6
see also content analysis, rigor,
validity
- comparison, 43, 47–58
logic of, 43, 47–51
paired, 43, 53–5
see also case studies
- conceptualization, 2, 4, 14, 15–19,
26–7, 28–9, 44–5
see also theories
- Consortium for Qualitative Research
Methods, 3, 212, 228, 233
- constitution, 50, 56, 76
see also causality, constructivism
- constructivism, 45, 48, 51–3, 57, 116,
121, 125, 168, 184, 188, 190–4
see also causality, constitution,
theories
- content analysis, 6, 151–67, 214, 223,
232
see also coding, discourse, pragmatic
analysis
- context, 4, 14, 16, 18, 54, 68, 71, 80,
83, 90, 116, 119, 124, 131–2, 134,
151, 160–3, 168, 185–6, 189
see also historiography,
interpretation
- contingency, 85, 86, 157
see also causality, context
- counter-factuals, 47, 120, 196
see also case studies, causality,
theories
- courses, 3, 7, 212, 213–14, 220, 224–6,
228–33
assignments, 224, 229–33
exercises, 225–6

- courses – *continued*
 syllabi, 3, 7, 212, 228–9
 see also pedagogy
 critical theory, 3, 28, 30, 45, 48, 57,
 214–15, 218–20
 see also post-colonial theory,
 post-modernism,
 post-structuralism, theories
 cultural competence, 5, 14, 63–5, 82,
 85–6, 89, 96, 210, 235, 236–7
 culture, 1, 13, 44, 62, 63–5, 72–3, 75,
 78, 80, 87–9, 90–1, 100, 107, 108,
 134, 151
 political, 88, 90–1, 134
 popular, 64, 72–3, 87–9, 90–1
 curb-cutting, 4, 38–41, 225
- data, 22–3, 24, 34–6, 50, 52, 69, 73,
 85–91, 108, 114, 116, 121, 151,
 152–5, 213–14, 232
 see also specific types of data gathering
 techniques
 deduction, 48, 49, 117, 190
 compare induction
 see also theories
 ‘deep hanging out’, 93, 99, 101–2,
 113, 225, 226, 230
 see also anthropology, ethnography
 delimitation, 20–1, 32, 63, 65–70,
 85–6, 90, 152–5, 185
 see also data, research questions,
 scope
 dictionary, *see* coding
 discourse, 1, 4–5, 14–15, 61–77, 78–92,
 107, 194, 212, 214, 223, 225, 233
 layering, 63, 66, 73–5, 90–1
 mapping, 70–3, 90–1
 see also genealogy, historiography,
 post-structuralism,
 representations
 dissertations, 22, 55, 63, 68, 78, 95
 see also funding (for research),
 research questions, writing
 Doty (Roxanne), 45, 78, 79
 doxa, 15, 17
 see also Bourdieu (Pierre)
 Duffy (Gavan), 84, 195, 213, 214, 223
 see also pragmatic analysis
- Dunn (Kevin), 14, 18, 31, 49, 51, 55,
 61, 72, 108, 119, 121, 124, 162,
 194, 213, 215, 216
 see also discourse, historiography,
 representations
- empiricism, 83, 92
 see also epistemology
 employment (in academia), *see* jobs (in
 academia)
 epistemology, 1–2, 5, 24–5, 28, 43, 51,
 58, 79–85, 91–2, 120–7, 195–9,
 213, 215, 228–9
 see also pluralism.
 ethics, 30, 31–2, 36, 38, 94, 109–13,
 124
 see also human subjects
 ethnography, 3, 5, 48, 63, 73, 93–113,
 198, 225, 230, 233
 see also anthropology, participant
 observation
 evidence, *see* data
 exercises, *see* courses
 experiments, 48, 53, 58, 184, 190,
 194, 198
- falsification, 1, 50
 see also positivism
 feminism, 28–42, 48, 212
 see also critical theory,
 epistemology, gender
 fetishizing (of methods), 215–17, 220,
 235
 fiction, *see* culture, data, texts
 field analysis, 15–17, 21, 236
 see also Bourdieu (Pierre)
 fieldwork, 96–9, 224
 see also anthropology, ethnography
 films, *see* culture, data, texts
 focus group, 221, 227, 234, 236
 see also interviews, surveys
 formal modeling, *see* modeling
 Foucault (Michel), 3, 5, 13, 15, 18, 45,
 48, 69, 71, 74, 84, 107, 212, 223,
 226
 compare Bourdieu (Pierre)
 see also discourse, genealogy,
 historiography,
 post-structuralism

- funding (for research), 1, 55, 107, 110, 112, 122, 217, 225, 234
see also jobs(in academia), research design, writing
- fuzzy sets, 56–7
see also case studies, comparison
- Geertz (Clifford), 93, 108
see also 'thick description'
- gender, 4, 11, 13, 15, 25, 29, 38, 48, 62, 66, 69, 71, 72, 89, 96, 97–8
see also feminism
- genealogy, 5, 50, 74–5, 113, 213
see also discourse, historiography
- Gramsci (Antonio), 13, 70
see also power
- grant-seeking, *see* funding(for research)
- Gusterson (Hugh), 18, 23, 29, 36, 46, 48, 63, 85, 89, 121, 225, 226, 235
see also anthropology, ethnography
- habitus*, 15, 17–18, 21–2, 78
see also Bourdieu (Pierre)
- Hermann (Margaret), 51, 84, 105, 115, 116, 214, 223
see also content analysis, personality
- hermeneutics, *see* interpretation, language, meaning
- heuristics, 7, 52, 73, 190, 195–6
- hierarchy, *see* power
- historiography, 1, 2, 3, 5, 53–4, 61, 78–92, 94, 116, 189, 232
 archives, 5, 87–8, 94
see also narrative, representations, social sciences
- Hoffmann (Matthew), 50, 53, 116, 211, 216, 217, 223, 234
see also agent-based modeling
- human subjects, 29, 30, 34, 109–13
see also ethics
- humanities, 12
see also anthropology, historiography, social sciences
- ideal type, 47, 56, 57, 61
see also typologies
- identity, 6, 50–1, 75, 78–92, 96, 105, 115, 116, 135–6
- images, *see* texts
- induction, 55, 190, 218
compare deduction
- inequality, *see* power
- interdisciplinary research, *see* social sciences
- international relations (as a discipline), 1, 3, 4–5, 7, 12, 21, 43, 44–5, 48, 51, 53, 201, 214, 216, 227–8
see also social sciences
- internet, *see* data, texts
- interpretation, 1–7, 78–85, 91, 190, 213, 229
see also language, meaning
- interviews, 3, 5, 29, 36, 69, 73, 78, 89, 94, 103–6, 116, 117–18, 119, 121, 153, 214, 220, 221, 225, 227, 236
see also data, participant observation, surveys
- IRB (Institutional Review Board), *see* human subjects
- jobs (in academia), 1, 25, 58, 201, 234
- King (Gary), Keohane (Robert) and Verba (Sidney), 47, 184, 214, 218
- Klotz (Audie), 72, 85, 116, 196, 219, 221, 229, 231
see also case studies
- Laclau (Ernesto) and Mouffe (Chantal), 67, 74
- language, 1, 2, 17, 61, 63, 78, 89, 220
see also culture, interpretation, meaning
- leadership, 82–3, 131–50, 151–67
see also agency(agents), personality
- Leander (Anna), 29, 44, 48, 65, 70, 79, 81, 82, 85, 90, 192, 194, 213, 235
see also Bourdieu (Pierre), conceptualization
- levels of analysis, *see* macro/micro distinction
- linguistics, 44, 80
see also language, meaning, pragmatic analysis, speech acts

- logic, 92, 180–3, 196, 198
see also pragmatic analysis
- 'long conversation', 14, 83–4
- macro-level of analysis, *see*
 historiography, process tracing,
 structure
- macro/micro distinction, 5, 20–2, 43,
 49, 82, 108, 115–16, 122–3, 184,
 187, 190, 197
see also agency(agents), structure
- magazines, *see* culture, data, texts
- meaning, 1, 2, 3, 14, 18, 63, 71, 76,
 80, 113, 151–2, 169, 186
see also discourse, language
- mechanisms, 2, 3, 5–6, 49, 50, 80,
 114, 115–16, 121, 122
see also causality
- metaphors, 62, 74, 107, 169, 177, 188
see also discourse, texts
- micro-level of analysis, *see* agency,
 process tracing
- modeling, 2, 183, 188–9, 195–9, 202,
 212, 216–17, 220
see also agent-based modeling,
 rational choice
- monuments, *see* texts
- music, *see* culture, data, texts
- narrative, 2, 5, 47, 49, 78, 82–3, 87,
 91, 98, 116, 169
see also causality, historiography,
 texts
- National Science Foundation, 109,
 216
see also funding(for research),
 research questions
- nationality, 63, 65, 66, 68, 96, 118
- network analysis, 217
- Neumann (Iver), 11, 15, 26, 51, 54,
 79, 81, 84, 86, 87, 89, 108, 124,
 162, 201, 213, 215, 235
see also discourse
- ontology, 1–2, 4, 5, 30, 44, 49, 115,
 200
- operationalization, 19–23
see also conceptualization
- Orientalism, *see* post-colonial theory
- participant observation, 5, 31, 36, 69,
 94, 99–103, 122, 220, 230
see also ethnography
- pedagogy, 3, 7, 11, 20, 38, 211–17,
 221–37
see also courses
- performance, 105–6
see also texts
- personality, 138–47, 148, 212, 223,
 226
see also psychiatry, psychology
- photographs, *see* texts
- plausibility probe, 51–2, 196
see also case studies, theories
- pluralism, 2–4, 7, 51, 120, 124, 202,
 212, 218–19, 222, 231, 235
see also epistemology;
 tools/toolbox/toolkit
- political science, 2, 3, 7, 88, 93, 108,
 127, 211–12, 217–18, 222, 227–8
see also social sciences
- positivism, 1–3, 50, 105, 108, 114,
 116, 120, 124, 125, 213, 229
compare critical theory,
 post-modernism,
 post-structuralism, relativism
see also epistemology, falsification
- Post (Jerrold), 115, 167, 194, 223, 226
see also personality, psychiatry,
 psychology
- post-colonial theory, 28, 30, 79, 81
see also critical theory,
 post-structuralism
- post-modernism, 1–3, 47, 49, 50,
 79–85
compare positivism, falsification
see also critical theory,
 epistemology,
 post-structuralism, relativism
- post-structuralism, 14, 28, 30, 70,
 79–127
see also critical theory,
 post-modernism
- power, 13, 15, 16, 18–19, 25, 28, 30,
 38, 48, 70, 80–2, 107, 217–19,
 223, 235
- practices, 11, 14, 15, 18–19, 78, 80–1
see also discourse

- pragmatic analysis, 168–86, 198, 223, 227
see also linguistics, speech acts, language, meaning
- press conferences, *see* data
- process tracing, 5–6, 15, 43, 49, 95, 114–227, 188, 190, 198, 202, 232
see also causality, mechanisms
- psychiatry, 6, 131–2, 134, 138–9, 140–7
 narcissist, 140–2
 obsessive-compulsive, 142–4
 paranoid, 144–7
see also personality, psychology
- psychology, 6, 117, 124, 131–50
 psychobiography, 131, 133, 134–7, 138, 148
see also psychiatry, personality
- publishing, *see* writing
- qualitative methods, definition of, 3, 4, 6, 43
- qualitative-quantitative distinction, 3, 43, 57, 151, 155–7, 188, 201, 211–20, 228, 233–5
- quantitative methods, 2, 3, 198, 213, 216–17
see also modeling, statistics
- race, 44–5, 46, 48, 51–2, 54, 55, 71, 72, 80, 89, 96
see also conceptualization, post-colonial theory
- rational choice, 2, 49, 76–7, 92, 116, 121, 124, 189, 190, 200
see also modeling, rationality
- rationality, 18, 44, 80, 82, 92, 113, 137, 142–3, 189
see also agency(agents), rational choice
- reflexivity, 4, 23–6, 28–42, 76, 77
see also interpretation
- relativism, 1, 91–2
see also post-modernism
- reliability, *see* coding
- representations, 5, 14, 25, 61–2, 63, 70–3, 74, 75, 78–92, 151, 155–7
see also discourse, historiography
- research design, 4, 28, 33–4, 53, 55, 213–15, 219–20, 229, 231
- research questions, 1, 4, 12–15, 29, 32, 46, 93, 109–13, 152, 223
 legitimacy of, 1, 29, 109–10
- rhetoric, *see* language, meaning
- rigor, 1, 2, 23–4, 62, 63, 69, 92, 94, 108, 113, 126, 168, 173, 184, 197, 200, 236–7
see also validity, epistemology
- rules, 14, 17, 108, 116, 191–2
- Said (Edward), *see* post-colonial theory
- sampling, 32–5, 98
see also data, interviews, surveys
- scientific realism, 116, 125–6, 197
see also epistemology
- scope, 20, 22–3, 42, 117, 183–6, 198
see also delimitation
- semiotics, *see* linguistics, language, meaning, pragmatic analysis, speech acts, texts
- signs, *see* linguistics, language, meaning, texts
- simulations, 6, 53, 188–90, 197–8, 199, 202–6
see also modeling
- social sciences, 1–3, 7, 12, 15, 24, 187, 198, 214, 218, 227
see also specific disciplines
- socialization, 6, 115, 116–20, 192
- sociology, 104, 222
see also social sciences
- speech acts, 2, 6, 168, 174–80, 184, 185–6, 195
see also linguistics, pragmatic analysis
- speeches, 162–3
see also data
- standards (for evaluating research), *see* rigor
- statistics, 2, 3, 47–8, 49, 116, 188, 212, 214, 215, 220, 232
see also quantitative methods
- structure, 70, 73–4, 78, 80, 83, 87, 114, 124, 200
compare agency(agents), post-structuralism

surveys, 69, 94, 104, 116, 215, 221,
235–7

see also data

syllabi, *see* courses

symbols, 13, 151

see also meaning, texts

tautology, 50

television, *see* culture, data, texts

texts, 3, 5, 6, 63, 64, 65–70, 72, 75, 86,

87, 89, 108, 180

genre, 69

monuments', 67, 70, 86

pioneer, 72

visual, 72, 75, 87, 89, 180

see also data, discourse

theories, 4, 11–12, 23, 27, 45, 47–8,

51, 52–3, 58, 93–4, 115, 116, 120,

184, 196, 198, 215–16

building, 52–3, 58.

testing, 47–8, 52, 58, 93–4, 116,

184, 198

'thick description', 22, 58, 78, 83,
108, 152

see also ethnography

tools/toolbox/toolkit, 2, 3, 7, 11–27,

75, 188, 202, 218, 219, 221, 236

travelogues, *see* culture, data, texts

triangulation, 119, 126, 227–8, 236

typologies, 56–7

see also ideal type, theories

unit of analysis, 43, 54, 151, 157–8

validity, 23–6, 83, 91–2, 119–20, 131,

151, 154–5, 164–6, 168, 173, 184,

189, 197–9, 213, 214, 219, 232

see also epistemology, rigor

variables, *see* theories

writing, 1, 22, 26, 29, 37, 44, 55, 64,

66–7, 68, 73, 76, 91, 94, 106–9,

112–13, 121, 135, 213, 223, 234

Introduction to quantitative and qualitative research

Dr Liz FitzGerald

Institute of Educational Technology

Research and research methods

- Research methods are split broadly into quantitative and qualitative methods
- Which you choose will depend on
 - your research questions
 - your underlying philosophy of research
 - your preferences and skills

Basic principles of research design

Four main features of research design, which are distinct, but closely related

- **Ontology:** How you, the researcher, view the world and the assumptions that you make about the nature of the world and of reality
- **Epistemology:** The assumptions that you make about the best way of investigating the world and about reality
- **Methodology:** The way that you group together your research techniques to make a coherent picture
- **Methods and techniques:** What you actually do in order to collect your data and carry out your investigations
- These principles will inform which methods you choose: you need to understand how they fit with your 'bigger picture' of the world, and how you choose to investigate it, to ensure that your work will be coherent and effective

Four main schools of ontology

(how we construct reality)

Ontology	Realism	Internal Realism	Relativism	Nominalism
Summary	The world is 'real', and science proceeds by examining and observing it	The world is real, but it is almost impossible to examine it directly	Scientific laws are basically created by people to fit their view of reality	Reality is entirely created by people, and there is no external 'truth'
Truth	There is a single truth	Truth exists, but is obscure	There are many truths	There is no truth
Facts	Facts exist, and can be revealed through experiments	Facts are concrete, but cannot always be revealed	Facts depend on the viewpoint of the observer	Facts are all human creations

However, none of these positions are absolutes.

They are on a continuum, with overlaps between them.

Epistemology

i.e. the way in which you choose to investigate the world

Two main schools are **positivism** and **social constructionism**:

- **Positivists** believe that the best way to investigate the world is through objective methods, such as observations. Positivism fits within a realist ontology.
- **Social constructionists** believe that reality does not exist by itself. Instead, it is constructed and given meaning by people. Their focus is therefore on feelings, beliefs and thoughts, and how people communicate these. Social constructionism fits better with a relativist ontology.

Methodology

- Epistemology and ontology will have implications for your methodology
- Realists tend to have positivist approach
→ tend to gather quantitative sources of data
- Relativists tend to have a social constructionist approach
→ tend to gather qualitative sources of data
- Remember these are not absolutes! People tend to work on a continuum → role for mixed methods and approaches
- Also consider the role of the researcher*: internal/external; involved or detached?

A note about data

- Quantitative data is about quantities, and therefore numbers
- Qualitative data is about the nature of the thing investigated, and tends to be words rather than numbers
- Difference between primary and secondary data sources
- Be aware of research data management practices and archives of data sets (both in terms of downloading and uploading)

Choosing your approach

- Your approach may be influenced by your colleagues' views, your organisation's approach, your supervisor's beliefs, and your own experience
- There is no right or wrong answer to choosing your research methods
- Whatever approach you choose for your research, you need to consider five questions:
 - What is the **unit of analysis**? For example, country, company or individual.
 - Are you relying on **universal theory or local knowledge**? i.e. will your results be generalisable, and produce universally applicable results, or are there local factors that will affect your results?
 - Will **theory or data come first**? Should you read the literature first, and then develop your theory, or will you gather your data and develop your theory from that? (N.B. this will likely be an iterative process)
 - Will your study be **cross-sectional or longitudinal**? Are you looking at one point in time, or changes over time?
 - Will you **verify or falsify a theory**? You cannot conclusively prove any theory; the best that you can do is find nothing that disproves it. It is therefore easier to formulate a theory that you can try to disprove, because you only need one 'wrong' answer to do so.

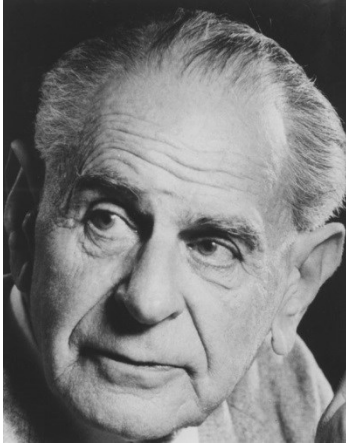
Quantitative approaches

- Attempts to explain phenomena by collecting and analysing numerical data
- Tells you if there is a “difference” but not necessarily why
- Data collected are always numerical and analysed using statistical methods
- Variables are controlled as much as possible (RCD as the gold standard) so we can eliminate interference and measure the effect of any change
- Randomisation to reduce subjective bias
- If there are no numbers involved, its not quantitative
- Some types of research lend themselves better to quant approaches than others

Quantitative data

- Data sources include
 - Surveys where there are a large number of respondents (esp where you have used a Likert scale)
 - Observations (counts of numbers and/or coding data into numbers)
 - Secondary data (government data; SATs scores etc)
- Analysis techniques include hypothesis testing, correlations and cluster analysis

Black swans and falsifiability



<https://www.flickr.com/photos/Iselibrary/IMAGELIBRARY/5>

- **Falsifiability** or refutability of a statement, hypothesis, or theory is the inherent possibility that it can be proven false
- Karl Popper and the black swan; deductive c.f. inductive reasoning

- Hypothesis testing
- Start with null hypothesis
i.e. H_0 – that there will be no difference



Type I and Type II errors

Never confuse Type I and II errors again:

Just remember that the Boy Who Cried Wolf caused both Type I & II errors, in that order.

First everyone believed there was a wolf, when there wasn't. Next they believed there was no wolf, when there was.

Substitute "effect" for "wolf" and you're done.

Kudos to @danolner for the thought. Illustration by Francis Barlow
"De pastoris puero et agricolis" (1687). Public Domain. Via [wikimedia.org](https://commons.wikimedia.org/wiki/File:De_pastoris_puero_et_agricolis.jpg)

Analysing quant data

- Always good to group and/or visualise the data initially → outliers/cleaning data
- What average are you looking for?
Mean, median or mode?
- Spread of data:
 - skewness/distribution
 - range, variance and standard deviation

What are you looking for?

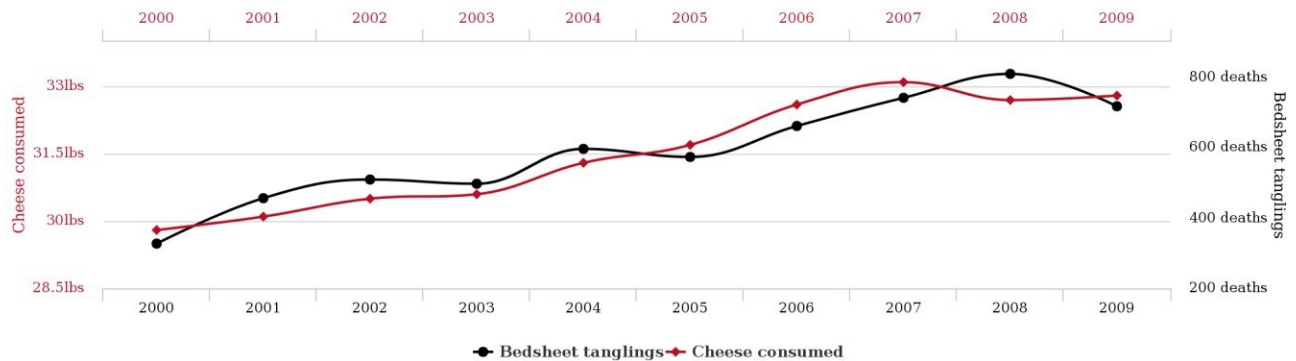
- Trying to find the signal from the noise
- Generally, either a **difference** (between/within groups) or a **correlation**
- Choosing the right test to use:
parametric vs **non-parametric** (depends what sort of data you have – interval/ratio vs nominal/ordinal and how it is distributed)
- Correlation does not imply causation!

Example correlations

Per capita cheese consumption

correlates with

Number of people who died by becoming tangled in their bedsheets



tylervigen.com

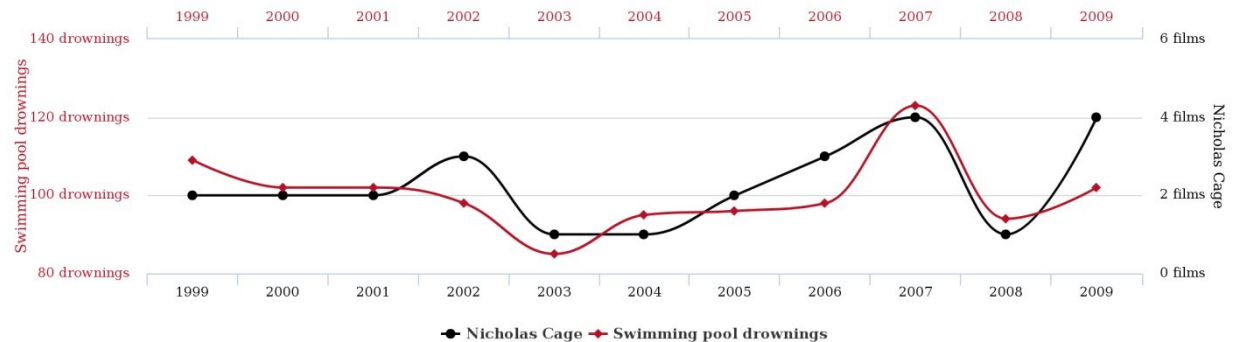
From 'Spurious correlations' website

<http://www.tylervigen.com/spurious-correlations>

Number of people who drowned by falling into a pool

correlates with

Films Nicolas Cage appeared in



tylervigen.com

Interpreting test statistics

- **Significance level** – a fixed probability of wrongly rejecting the null hypothesis H_0 , if it is in fact true. Usually set to 0.05 (5%).
- **p value** - probability of getting a value of the test statistic as extreme as or more extreme than that observed by chance alone, if the null hypothesis H_0 , is true.
- **Power** – ability to detect a difference if there is one
- **Effect size** – numerical way of expressing the strength or magnitude of a reported relationship, be it causal or not

Example of quant data/analysis*

- Matched users were those who learning styles were matched with the lesson plan e.g. sequential users with a sequential lesson plan. Mismatched participants used a lesson plan that was not matched to their learning style, e.g. sequential users with a global lesson plan.
- H_0 – there will be no statistically significant difference in knowledge gained between users from different experimental groups
- H_1 – students who learn in a matched environment will learn significantly better than those who are in mismatched environment
- H_2 – students who learn in a mismatched environment will learn significantly worse than those who learn in a matched environment

Interpreting test statistics

TABLE 6.3: Difference in knowledge gained between matched and mismatched groups

Group type	N	Avg. knowledge gained (points)	Standard deviation
Matched	39	1.23	3.29
Mismatched	43	1.98	3.65

- Statistical testing was carried out using a univariate ANOVA in SPSS, to determine if there was any significant difference in knowledge gained.
- Initial conjecture suggests that the mismatched group actually performed better than the matched group.
- However, the difference between the two groups was not significant ($F(1,80)=0.939, p=0.34, \text{partial eta squared} = 0.012$) and hence hypotheses 1 and 2 can be rejected.

What quant researchers worry about

- Is my sample size big enough?
- Have I used the correct statistical test?
- have I reduced the likelihood of making Type I and/or Type II errors?
- Are my results generalisable?
- Are my results/methods/results reproducible?
- Am I measuring things the right way?

What's wrong with quant research?

- Some things can't be measured – or measured accurately
- Doesn't tell you **why**
- Can be impersonal – no engagement with human behaviours or individuals
- Data can be static – snapshots of a point in time
- Can tell a version of the truth (or a lie?)
“Lies, damned lies and statistics” – persuasive power of numbers

Qualitative approaches

- Any research that doesn't involve numerical data
- Instead uses words, pictures, photos, videos, audio recordings. Field notes, generalities. Peoples' own words.
- Tends to start with a broad question rather than a specific hypothesis
- Develop theory rather than start with one
→ inductive rather than deductive

Gathering qual data

- Tends to yield rich data to explore **how** and **why** things happened
- Don't need large sample sizes (in comparison to quantitative research)
- Some issues may arise, such as
 - Respondents providing inaccurate or false information – or saying what they think the researcher wants to hear
 - Ethical issues may be more problematic as the researcher is usually closer to participants
 - Researcher objectivity may be more difficult to achieve

Sources of qual data

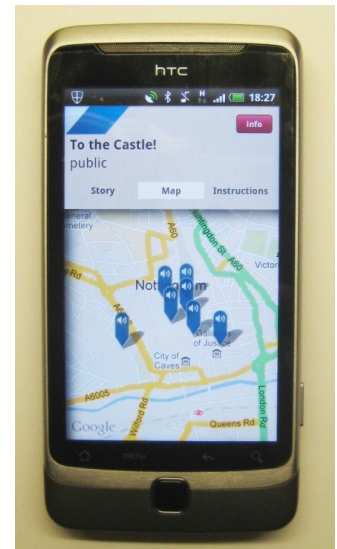
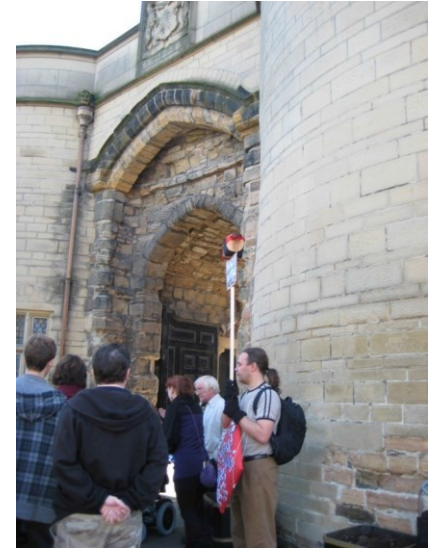
- Interviews (structured, semi-structured or unstructured)
- Focus groups
- Questionnaires or surveys
- Secondary data, including diaries, self-reporting, written accounts of past events/archive data and company reports;
- Direct observations – may also be recorded (video/audio)
- Ethnography

Analysing qual data

- Content analysis
- Grounded analysis
- Social network analysis (can also be quant)
- Discourse analysis
- Narrative analysis
- Conversation analysis

Example of qual data research*

- Describing and comparing two types of audio guides: person-led and technology-led
- Geolocated audio to enable public, informal learning of historical events
- Data sources: questionnaires, researcher observations, and small focus groups



Data analysis and findings

- Comparison of the two different walks
 - Differences/similarities of the walks
 - Issues surrounding participant engagement
- Thematic analysis
 - Mode of delivery
 - Number of participants and social interactions
 - Geographical affordances of places and locations
 - User experience
 - Opportunities for learning
 - Other factors
- Findings, lessons learned, recommendations

What qual researchers worry about

- Have I coded my data correctly?
- Have I managed to capture the situation in a realistic manner?
- Have I described the context in sufficient detail?
- Have I managed to see the world through the eyes of my participants?
- Is my approach flexible and able to change?

What's wrong with qual research?

- It can be very subjective
- It can't always be repeated
- It can't always be generalisable
- It can't always give you definite answers in the way that quantitative research can
- It can be easier to carry out (or hide) 'bad' (poor quality) qual research than 'bad' quant research

Other aspects of research design

- Validity
- Reliability
- Trustworthiness*
 - **Dependability**: showing that the findings are consistent and could be repeated
 - **Confirmability**: a degree of neutrality or the extent to which the findings of a study are shaped by the respondents and not researcher bias, motivation, or interest
 - **Credibility**: confidence in the 'truth' of the findings
 - **Transferability**: showing that the findings have applicability in other contexts

Summary

- The type of approach you choose will be determined by your research question, your epistemological and ontological stances and your skills or ability to utilise a certain approach
- For most people in ed tech, a mixed methods approach will be used
- So long as you make an informed choice and can justify it, it should be fine 😊
- Just be aware of the limitations of your approach(es) and try to compensate where necessary

Acknowledgments and further links

- Some content borrowed from SkillsYouNeed website
(<http://www.skillsyouneed.com/learn/research-methods.html>)

Other useful links:

- Introduction to Quantitative and Qualitative Research Models (William Bardebes). PDF at <http://tinyurl.com/qq-models>
- Methods Map: <http://www.methodsmap.org>
- Ready To Research: <http://readytoresearch.ac.uk>
- Methods@Manchester:
<http://www.methods.manchester.ac.uk/resources/categories>
- Research Data Management training: <http://datalib.edina.ac.uk/mantra/>

INTRoDucTIoN TO

SOCIAL RESEARCH

QUANTITATIVE & QUALITATIVE APPROACHES

KEITH F PUNCH

ISAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

List of tables	xiii
List of figures	xv
Preface	xvii
Companion website	xix
About the author	xxi
 1 Introduction	 ■
2 Theory and Method in Social Science Research	13
3 Ethics in Social Science Research	35
4 Research Questions	57
5 From Research Questions to Data	73
6 Literature Searching and Reviewing	93
7 Qualitative Research Design	113
8 Collecting Qualitative Data	143
9 Analysing Qualitative Data	167
10 Quantitative Research Design	205
11 Collecting Quantitative Data	227
12 Analysing Quantitative Data	251
13 The Internet and Research	279
14 Mixed Methods and Evaluation	301
15 Research Writing	327
Glossary	347
Appendix 1	355
References	359
Index	377

ten

QUANTITATIVE RESEARCH DESIGN

Contents

- 10. Research design
- 10. Some background
- 10. Independent, dependent and control variables
- 10 The experiment
- 10 Quasi-experimental and non-experimental design
- 10 Relationships between variables: the correlational survey
- 10 Relationships between variables: causation and accounting for variance
- 10 Multiple linear regression (MLR) as a general strategy and design
- 10 Controlling variables
- Chapter summary
- Key terms
- Exercises and study questions
- Further reading
- Notes

Learning Objectives

After studying this chapter you should be able to:

- Describe the similarities and differences between comparing groups and relating variables, as strategies in quantitative research
- Define independent, dependent and control variables
- Describe the basic characteristics of an experiment
- Show how the logic of experimental design extends to quasi-experimental and correlational survey designs
- Explain the key concept of accounting for variance
- Explain how multiple linear regression fits in with accounting for variance

In the most general terms, quantitative research does three main things:

- it conceptualises reality in terms of variables;
- it measures these variables; and
- it studies relationships between these variables.

Thus variables (and variance) are the central concepts in quantitative research.

Chapter 11 will deal with variables and their measurement. This chapter focuses on relationships between variables. From a quantitative design point of view, we can study relationships between variables either by comparing groups, or by relating variables directly. One theme of this chapter is therefore the broad division in the logic of quantitative design between comparing groups, on the one hand, and relating variables, on the other. We can see this by looking briefly at some methodological history in Section 10.2. Three main types of design follow from this broad division – experiments, quasi-experiments and correlational surveys. A second theme of the chapter is the shift from comparison-between-groups to relationships-between-variables, as a way of thinking, and to regression analysis as a strategy and design for implementing this shift. Running through both themes are the ideas of independent, control and dependent variables. We begin the chapter by reviewing the concept of research design, described in Chapter 7.

Research design

In Chapter 7 research design was described as the overall plan for a piece of research, including four main ideas – the strategy, the conceptual framework, the question of who or what will be studied, and the tools to be used for collecting and analysing data. Together, these four components of research design situate the researcher in the empirical world. Design sits between the research questions and the data, showing how the research questions will be

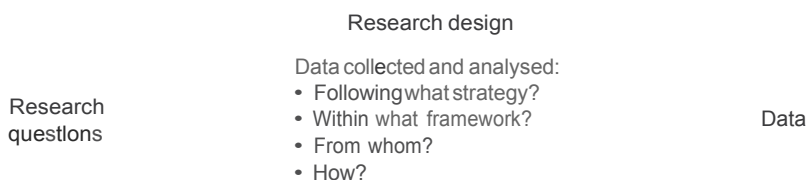


FIGURE 10 Research design connects research questions to data

connected to the data, and what tools and procedures to use in answering them. Therefore it needs to follow from the question and fit in with the data.

Thus, as in Chapter 7 and as shown again in Figure 10.1, in considering design we are moving from what data will be needed to answer the research questions (the empirical criterion, see Section 5.1) to how and from whom the data will be collected. The same four questions are used: the data will be collected:

Following what strategy?
Within what framework?
From whom?
How?

In quantitative studies, where variables are central, the design and conceptual framework tend to come together. The design shows how the variables are arranged, conceptually, in relation to each other. In other words, it shows diagrammatically the strategy behind the research. As stressed in Chapter 7, all research design is driven by strategy. The conceptual framework also shows the structure of the proposed study in terms of its variables. While quantitative research design tends to fall towards the tightly structured end of the structuring continuum, it varies in how much the situation is contrived for research purposes, as this chapter will show.

Some background

A brief sketch of some of the methodological history of quantitative research provides background both for this chapter and for Chapter 12, on quantitative data analysis.

Empirica! social science research, as we know it today, began some 150 years ago (with the exception of economics, which has a much longer history). The early social scientists, especially in psychology and sociology, were impressed by the progress of the natural sciences, especially physics and chemistry, and set out to imitate their use of the scientific method in building knowledge. They saw the core of the scientific method as two things – the experiment and measurement. We describe the experiment later, but its central idea involves the artificial manipulation of some treatment variable(s) for research purposes, setting up controlled comparison

groups. In the simplest case, the comparison groups are alike in all respects – that is alike on all other variables – except for their differential exposure to the treatment variable. The other variables are controlled by the design. The aim is to relate the treatment variable(s) to outcome variable(s), having controlled the effects of other variables. The experiment was seen as the basis for establishing cause-effect relationships between variables, and its outcome (and control) variables had to be measured. Thus most early social science research was characterised by experimental design and by measurement.

In the 1950s and 1960s, quantitative researchers in social science began to broaden the scope of the experiment, partly because of its limitations. There was no questioning the logic of the experiment, but limitations to its applicability, both practical and ethical, forced this development. The logic of the experiment was extended first to quasi-experimental and then to non-experimental situations. These terms are explained in Section 10.5. This happened because many of the most important questions in social science research could not be studied by experimental design. Yet there were many examples of naturally occurring treatment groups (see Section 10.5), where the comparisons of interest were possible, but where they had not been set up specifically for research purposes. The development was therefore to apply the principles of experimental design to these quasi-experimental situations, studying these naturally occurring treatment groups. Since these comparison groups had not been set up for research, other (extraneous) variables were not controlled in the design. Therefore, it was necessary to develop techniques for controlling extraneous variables in the analysis of data, since, with the impossibility of true experimentation, they could not be controlled in the design. Put simply, what was developed was a statistical approximation to the desired experimental situation where the comparison groups were alike in all respects, on these other variables. This was done through the statistical control of extraneous variables in data analysis, rather than through the physical control of these variables in the design. These ideas are more fully described in Sections 10.4 through to 10.9. In these developments, measurement continued to be central – the introduction of more variables only accentuated the need for measurement.

These developments led to two main strands within the fields of quantitative design and data analysis:

First, the comparison-between-groups strand, based on the experiment, and with the t-test and analysis of variance as its main statistical features;

Second, the relationships-between-variables strand, based on non-experimental reasoning, with correlation and regression as its main features. I will call this second strand the correlational survey strand.

Comparing the direction of thinking behind these two strands is interesting. The true experiment looks 'downwards' or 'forwards', as it were, from the independent variable to the dependent variable, or from causes to effects. The central question here is: What is the effect of this cause? On the other hand, the correlational survey looks 'upwards' or 'backwards', from dependent variable to the

independent variable, from effects to causes. The central question here is: What are the causes of this effect? Because this latter approach takes the world as a given, studying it after things have happened, it is sometimes called *ex post facto* research – that is, research that occurs after the fact. Mapping the variance in the dependent variable, and *accounting for variance in the dependent variable*, become two central notions in this way of thinking, and they are important themes in this and the next two chapters.

The above description is most typical of applied social science areas, especially those with a sociological bias, including education. The two strands developed in a different way in psychology and educational psychology, as Cronbach (1957) has pointed out. However, the end result is much the same. Cronbach called the two strands the 'experimentalists' and the 'correlationists'. The experimentalists create variation in the treatment variable, in order to study the consequences of doing this. They study how nature is put together, not through taking nature as it is, but through changing it and understanding the consequences of these changes. The correlationists, on the other hand, study the natural correlations occurring in nature. There is no manipulation to introduce changes, but rather the study of nature as it is (Shulman, 1988).

The two strands, the comparison-between-groups strand and the relationship-between-variables strand, are related to each other, particularly when it comes to the analysis of data. But they are also important distinct emphases, and a convenient way to present the material of this chapter. We will deal with the experiment first, and then move through the quasi-experiments to the correlational survey. This is because it is important to understand the logic of experimental design, and the developments which have flowed from it. Before that, however, we have to deal with some terminology.

Independent dependent and control variables

Discussing causation in Chapter 5, I pointed out that technical research language avoids the use of the terms 'cause' and 'effect'. The most common terms substituted, and those that will mostly be used here, are *independent* variable (for cause) and *dependent* variable (for effect). However, they are not the only terms used, as shown in Table 5.1. In experimental design, common terms are also 'treatment' and 'outcome' variable, and the treatment variable is often also called the experimental variable. But 'independent' and 'dependent' variable are the most widespread terms, and apply in both the experimental and non-experimental (survey) situations. In addition to independent and dependent variables, we need now to introduce the idea of control variables.

A *control variable* is a variable whose effects we want to remove or control. We want to control this variable because we suspect that it might confound, in some way, comparisons we want to make or relationships we want to study. It is extraneous to the variables that we really want to study, but, at the same time, may influence these variables and the relationship between them (Rosenberg, 1968).

Therefore we want to remove its effects. Technical synonyms for 'remove its effects' are 'partial it out' or 'control it'. In addition, the term 'covariate(s)' is often used as a synonym for control variable(s), and analysis of covariance is the most general of the techniques for controlling variables. A description of ways of controlling these extraneous variables is given later in this chapter (Section 10.9) and in Chapter 12. For the moment, we are thinking only of the conceptual role of control variables in a research design.

Now we have three general categories or types of variables:

independent variable(s)	control variable(s) (covariates)	dependent variable(s)
----------------------------	--	--------------------------

This shows the conceptual status of different variables in our thinking about research design. It is a general conceptual framework, showing the structure of a study in terms of these variables. The conceptual status for any variable may change from study to study, or from part to part within the one study. Thus, any particular variable may be an independent variable in one study, a dependent variable in another, and a control variable in a third. The researcher must of course make clear the conceptual status of each variable at each stage of the study.

The experiment

As noted in Section 10.2, one main strand in quantitative research design is the comparison-between-groups strand. The clearest case of this is the experiment. In research design, 'experiment' is a technical term with a precise meaning, which is made clear shortly. In discussing the logic of the experiment, we will use the simplest possible case of only two comparison groups.

The basic idea of an experiment, in social science research, is that two comparison groups are set up. Then we, as researchers, will do something (administer a treatment or manipulate an independent variable) to one of the groups. We call this group the experimental group or the treatment group. We do something different, or nothing at all, to the other group (we call this group the control group). We then compare the groups on some outcome or dependent variable. Our intention is to say that any differences we find in the outcome variable between the groups are due to (or caused by) the treatment or independent variable. In technical terms, we aim to attribute dependent (or outcome) variable differences between the groups to independent (or treatment) variable differences between the groups. This attribution is based on the important assumption that the groups are alike in all other respects. We will discuss this assumption shortly.

The experiment is based on comparisons between the groups. In the simplest case described above, the objective is to have the two groups alike in all respects, except that they receive different treatments – they have differential exposure to the independent variable. We then test for differences between them in the

outcome (dependent) variable. If the only difference between the groups is in the treatment they received (that is, in their exposure to the independent variable), then, because the independent variable occurs before the dependent variable, we have the strongest possible basis for inferring that differences in the dependent variable are caused by the independent variable. This is why the experiment has traditionally been the preferred design among so many quantitative researchers, especially in educational psychology. Box 10.1 shows the conceptual equivalence of terms used in this description of the experiment.

BOX 10.1

Terms that are Conceptually Equivalent

These terms are conceptually equivalent:

- administer a treatment to the experimental group
- manipulate the independent (or treatment) variable
- differential exposure to the independent variable.

Thus we manipulate an independent variable by administering treatment to an experimental group, which provides differential exposure to the independent variable.

The alike-in-all-respects criterion is the important assumption referred to earlier. How can this be achieved? How can the comparison groups be set up to be identical, except for differential exposure to the independent variable? Not easily, and historically different methods have been tried for achieving this. At one time, matching was favoured, whereby there was a deliberate effort to match the group members, one by one, in terms of relevant characteristics. However, one does not need many characteristics before this turns out to be impractical. Modern experimental design favours the random assignment of participants to comparison groups, as the way of meeting the alike-in-all-respects criterion.

This solution demonstrates a fundamental principle of quantitative reasoning. Random assignment of participants to treatment (or comparison) groups does not guarantee likeness or equality between the comparison groups. Rather, it maximises the probability that they will not differ in any systematic way. It is an ingenious way to control for the many extraneous variables that could differ between the groups, and therefore could invalidate conclusions about relationships between the independent and dependent variables based on comparisons between the groups. Random assignment of participants to treatment groups is a form of physical control of extraneous variables. When physical control of these variables by random assignment of participants to treatment groups is not possible,

researchers will resort to the statistical control of variables. This is where true experimental design gets modified into various quasi-experimental designs. This is described in Section 10.5.

To summarise, then, we have a true experiment if there is:

the manipulation of one or more independent variables for the purposes of the research; and
the random assignment of participants to comparison groups.

This description gives the essential logic of the experiment, but it is only an introduction to the topic of experimental design. Because situations in the real world are so variable, and because extraneous variables can influence experimental results in so many different ways, it has been necessary to modify and embellish this basic experimental design a great deal (see, for example, Kirk, 1995). Thus a wide variety of experimental designs has been developed, in order to ensure greater internal validity in different sorts of social science research situations. 'To ensure internal validity' here means to ensure better control of extraneous variables, or to eliminate rival hypotheses to the proposed causal one linking the independent and dependent variables. As a result of these developments, experimental design is a specialised topic in its own right. An important reference for the topic is the classic work by Campbell and Stanley (1963), in which they list the most common designs and the threats to the internal validity of those designs. Example 10.1 shows a number of research experiments.

EXAMPLE 10.1

Examples of Experiments

In 'Opinions and social pressure', Asch's (1955) classic experiment on compliance, male undergraduate students were recruited for a study of visual perception. Seven subjects were shown a large card with a vertical line on it and then asked to indicate which of three lines on a second card matched the original. Six of the group were accomplices of the researcher and gave false responses. The one 'real' subject was exposed to the subtle pressure of the other participants who presented a unanimous verdict.

Sherif et al. (1961) conducted a classic field experiment, *Intergroup Conflict and Cooperation: The Robber's Cave Experiment*, in which preadolescent American boys were brought into a summer camp in order to control and study the relations that developed among them.

In Project STAR (the Tennessee Student Teacher Achievement Ration experiment), Finn and Achilles (1990) studied the effect of reductions in class size on student academic achievement. Students were randomly assigned to classes of different sizes in 80 elementary schools.

Williams (1986) used the introduction of television into a remote Canadian community in the 1970s to study the effect of TV on children's cognitive skills.

Where it is possible to experiment, this design clearly provides the strongest basis for inferring causal relationships between variables. However, there are two problems that severely limit the applicability of the experiment in social science research. The first is practicality. It is simply not possible to investigate experimentally many of the questions of real interest and importance. Even with substantial funding, these questions remain out of reach, just on practical grounds. The second is ethics. Very often, questions of research interest are beyond the reach of the experiment, for a variety of ethical reasons.

However, despite these limitations, it is often still possible to make many of the comparisons we want to make, even if they are not set in a tight experimental design. There are situations where the comparisons that we want to make (and that we would have structured into an experiment were that possible) occur 'naturally', in the sense of not having been set up artificially for research purposes. These are called 'naturally occurring treatment groups'. How can we capitalise on these situations for research purposes? This question leads us to consider first quasi-experimental and then non-experimental designs. Both involve the extension of experimental reasoning to the non-experimental situation.

Quasi-experimental and non-experimental design

We can summarise the essential ideas here as follows:

In the quasi-experiment, comparisons are possible because of naturally occurring treatment groups. These naturally occurring treatment groups are fairly clear-cut, though not set up for research purposes. Therefore the experimental treatment is not controlled by the researcher, but the researcher has some control over when to measure outcome variables in relation to exposure to the independent variable. Some quasi-experiments are shown in Example 10.2. In the non-experiment, because the comparison groups as such are either not at all clear-cut or non-existent, the concept of naturally occurring treatment groups is broadened to naturally occurring variation in the independent variable. The researcher has little control over when to measure outcome variables in relation to exposure to the independent variable. The non-experiment is really now equivalent to the correlational survey.

EXAMPLE 10.2

Examples of Quasi-Experiments

In 'Comparison of feminist and non-feminist women's reactions to variants of non-sexist and feminist counselling', Enns and Hackett (1990) addressed the issue of matching client and counsellor interests along the dimension of attitudes towards feminism. The hypothesis tested was that feminist subjects would be more receptive

(Continued)

(Continued)

to a radical feminist counsellor whereas non-feminist subjects would rate the non-sexist and Liberal feminist counsellor more positively.

Glass's (1988) 'Quasi-experiments: the case of interrupted time series' described a number of quasi-experiments utilising time-series designs across several research areas: psychotherapy, road traffic accidents and fatalities, the stock market, self-esteem, anxiety, crime statistics and state school enrolments.

In *Experimental and Quasi-Experimental Designs for Research*, Campbell and Stanley (1963) described the formal characteristics, and the strengths and weaknesses, of ten different types of quasi-experimental designs.

In *Big School, Small School*, Barker and Gump (1964) studied the effects of school size on the lives of high school students and their behaviour, using samples of US schools of different sizes.

Shadish and Lueken (2006) report several education research examples using quasi-experiments with slightly different designs.

Thus there is a continuum of quantitative research designs here, where the true experiment is at the left-hand end, the non-experiment at the right-hand end, and the quasi-experiment in between. This continuum, shown in Figure 10.2, is about two things:

The researcher's ability to control exposure to the independent variable, and therefore how clear-cut the comparison groups are. In the experiment, the researcher manipulates the independent variable, and has control over the groups' exposure to it. In the quasi-experiment and the non-experiment, the researcher has no such control.

The researcher's ability to control when to take measurements on the dependent variable(s) in relation to exposure to the independent variable. Again, in the experiment the researcher can control this, taking dependent variable measurements at the most appropriate time. In the non-experiment, there is little opportunity to control this.

Thus, in both cases, researcher control is high at the left-hand end of this continuum, and low at the right-hand end.

We want to take advantage of naturally occurring treatment groups in a research situation. They provide the comparisons we want. But there is a logical difficulty in doing this, a clear threat to internal validity. It relates to the alike-in-all-respects criterion of the experiment. We may well find exactly the comparisons we want, in naturally occurring treatment groups, and we can certainly make the comparisons between these groups, with respect to one or more dependent (outcome) variables. But how can we be sure that there are not other differences between these naturally occurring comparison groups, over and above their differential exposure to the independent variable – differences that may themselves be responsible for any differences between the groups on dependent (outcome) variables? We have not been able to assign people randomly to these groups, to control variables physically, through the design. Therefore there is the real possibility of extraneous variable

Experiment	Quasi-experiment	Non-experiment (correlational survey)
Manipulation of independent variable(s)	Naturally occurring treatment groups	Naturally occurring variation in independent variable(s)
Random assignment to treatment groups	Statistical control of covariate(s)	Statistical control of covariate(s)

FIGURE 12.1 Spectrum of quantitative research designs

influences—that is, of systematic differences between the groups, on factors relevant to the dependent (outcome) variable.

The strategy to deal with this problem is to remove the influence of such possible extraneous variables by identifying them, measuring them and extracting their effects statistically. We control them statistically, in the analysis, using the rationale shown in Chapter 12 (Sections 12.3.3 and 12.4.7). Logically, controlling variables in this way achieves a statistical approximation to the desired physical situation of the experiment, where the comparison groups are alike in all respects except for their differential exposure to the independent variable. These extraneous factors become the control variables, or the covariates mentioned earlier. A covariate is thus an extraneous variable which is likely to be related to the outcome variable, and to differ between the comparison groups. The analysis of covariance (ANCOVA) is the technical name given to the statistical technique for controlling covariates. Control variable analysis is the more general term for the statistical control of extraneous variables.

Control variable analysis, and covariance analysis in particular, is an important and widely used quantitative research strategy and design. It applies when there are one or more extraneous variables whose effects we want to remove, in order to get a clearer picture of relationships between independent and dependent variables. All control variables have to be identified and measured before the implementation of the treatment(s). We cannot control a variable, or co-vary out its effects, during the analysis of data unless we have measurements on this variable. And we won't have measurements on it unless we have anticipated its possible effects, and designed its measurement into the study. This is another example of the benefits of the careful question development work recommended in Chapters 4 and 5.

Random assignment of participants to treatment groups, as in the true experiment, is the strongest design for demonstrating causality. But, given the great difficulty of doing this in real world research, control variable analysis in general and covariance analysis in particular are valuable in many research situations. It is therefore a major concept in quantitative design and analysis. It will come up again in Chapter 12, but its essential logic can be expressed in a few sentences:

To co-vary out one or more variables from a comparison between groups is to reach a statistical approximation to the (desired) physical situation where the groups are the same on the covariate(s). If they are the same on the covariates, the covariates cannot be responsible for differences on the outcome variables. Therefore outcome variable differences are more likely to be due to independent variable differences.

I have put this in terms of comparison-between-groups, in order to see it clearly. It applies just as well in studying relationships-between-variables. This is taken up again in Section 10.9, on the physical and statistical control of variables.

What sorts of variables should be seen as covariates? As always, logical considerations prevail. Following the above description, a variable should be controlled or co-varied out if it:

- is known or suspected to differ between the comparison groups; and
- is related to either the independent variable, or, more importantly, the dependent variable.

As Chapter 12 will show, the logic of the statistical technique of covariance analysis is to extract from the dependent variable the variance it holds in common with the covariate(s), and then to see if the remaining variance in the dependent variable is related to the independent variable. Thus, covariance analysis, like everything else in quantitative design and analysis, works on relationships between variables. It is time now to deal directly with this theme. This means we move from the first main strand in quantitative design (the comparison-between-groups strand) to the second main strand (the relationships-between-variables strand).

Relationships between variables: the correlational survey

In the quasi-experiment, treatment groups are reasonably clear-cut. In the non-experiment – that is, the correlational survey – we move from discrete comparison groups to naturally occurring variation in the independent variable. Instead of talking about separate comparison groups who differ on some variable of interest, we are now talking about a whole range of naturally occurring differences on this same variable. Discrete comparison groups, whether two or more, are simply a special case of this more general situation.¹ From now on, I will use the term 'correlational survey' instead of 'non-experiment' for this research design.

The word 'survey' has different meanings. It is sometimes used to describe any research that collects data (quantitative or qualitative) from a sample of people. Another meaning, common in everyday language, is a simple descriptive study, usually concerned with individual pieces of information, which are studied one piece at a time. Variables as such may not be involved, and continuous variables, as will be described in Chapter 11, are unlikely. This is sometimes called a 'status survey' or a 'normative survey' or a 'descriptive survey', and its purpose is mainly to describe some sample in terms of simple proportions and percentages of people who respond in this way or that to different questions. Such surveys are common today, especially in market research and political research.

The term 'correlational survey' is used here to stress the study of relationships between variables, and some surveys of this type are shown in Example 10.3. These relationships are often studied using conceptual frameworks similar to those used in experimental design. Thus, in this sort of survey, we conceptualise different variables as independent, control (or covariate) and dependent, as shown earlier. This illustrates

the point already made – the logic behind correlational surveys is based on the logic behind experimental design. Because we can only rarely experiment, research methodologists have applied the principles of experimental reasoning to the non-experimental research situation, developing logically equivalent non-experimental designs for those situations where variation occurs in the independent variables of interest, but where it is not possible to manipulate or control that variation for research purposes.² For this reason, it is important for researchers to understand the basic principles of experimental design, even if they are unlikely to design and use experiments.

EXAMPLE 10.3

Correlational Surveys

Bean and Creswell's (1980) study 'Student attrition among women at a Liberal arts college' investigated factors affecting student dropout rates at a small religious coeducational Liberal arts college in a Midwestern American city.

Blau and Duncan's (1967) influential book *The American Occupational Structure* looked at the movement from 'particularism' and 'ascription' to 'universalism' and 'achievement' by surveying occupational mobility in American society. The book included considerable material on the role of education in the intergenerational transmission of inequality. This was one of the first studies to use path analysis.

Equality of Educational Opportunity (Coleman et al., 1966) undertook the most comprehensive survey of the US school system, focusing mainly on the relationship between school characteristics and student achievement. Using a series of regression analyses, it produced the finding that school characteristics had little effect on student achievement. This led to the controversial conclusion that family background was more important than school characteristics in explaining differential achievement.

Peaker (1971) in his report, *The Plowden Children Four Years Later*, described a follow-up national survey of 3000 school-age children in the United Kingdom. Combined evidence from home and school from 1964 and 1968 is analysed and displayed.

Fifteen Thousand Hours: Secondary Schools and their Effects on Children (Rutter et al., 1979) is a large-scale study of 12 London secondary schools carried out over a three-year period. The study investigated whether schools and teachers have any effect on the development of children in their care.

We will now look at the relationships-between-variables strand of quantitative design, and see how that can be developed into an accounting-for-variance research design strategy.

Relationships between variables: accounting for variance

To say that two variables are related is to say that they vary together, or co-vary, or share common variance. What variance means, and the different

ways in which variables can vary together, will be explored in Chapters 11 and 12, but the essential idea of covariance is that the variables hold some of their variance in common. When two variables hold some of their variance in common, we can use the concept of accounting for variance and say that one variable accounts for (some of the) variance in the other. We can also say that one variable explains some of the variance in another, but accounting for variance is the more common description.

In Chapter 5 we took a brief philosophical look at the concept of causation single and multiple. We saw how important this concept is in science, since we want to find the cause(s) of events or effects. But we saw that we cannot do this directly, because of the metaphysical element in the concept of causation. Therefore we may start with ideas of causation in mind, but we need to 'translate' these ideas to make them suitable for empirical research, rephrasing our research questions to replace causal language.

One way to do this is to change questions of the form 'What causes Y?' into 'What causes Y to vary?' and then into 'How can we account for variance in Y?' The first rephrasing introduces the term 'vary'. To vary means to show differences, so now we are looking for and focusing on differences in Y, on variance in Y. This is important – our strategy of inquiry in order to learn about Y is to look for variance in Y, for differences in Y. This simple conceptual step is fundamental to a great deal of empirical inquiry, and underlines the key importance of the concept of variance in research. This same point comes up again in different ways: ... measurement in Chapter 11 (see especially Section 11.8) and in data analysis in Chapter 12 (Section 12.4). Now we almost have a form of the question that we can operationalise, but we still have to get rid of the troublesome word 'cause'. So we rephrase a second time, and now the question is in terms of accounting for variance.

Thus accounting for variance becomes a crucial step in the way we proceed in empirical research, especially *ex post facto* research. Variance means differences – this is why it is often said that the scientific method works on studying differences. A main strategy of empirical science is to find out how some dependent variable of interest varies, and then to account for this variance. The idea of learning about a phenomenon by studying its variation and accounting for this variation applies also in qualitative research, as can be seen in some approaches to the analysis of qualitative data, and especially in grounded theory (see Chapter 9).

Returning to quantitative research, we now have a research strategy we can operationalise. This is because, if two variables are related, one can be said to account for (some of the) variance in the other. This is the crux of the matter. The way we account for variance in a dependent variable is to find the independent variables to which it is related.

As pointed out in Chapter 5, we have moved well past simple one-variable causation, accepting the idea of multiple causation for any particular dependent variable. This is the commonly occurring design shown in the top right-hand cell of Figure 5.1 in Chapter 5. We accept that several (maybe many) factors will be necessary to give us a full causal picture for this dependent variable. In the language of this chapter, we have several independent variables and one dependent variable. If

we can account for most of the variance in our dependent variable with a particular set of independent variables, and if we know the importance of each of the independent variables in accounting for this dependent variable variance, then we understand the dependent variable very well – how it varies and how to account for its variance. Just as important, we would also have clear indications about which independent variables to concentrate on, in order to bring about changes in the dependent variable.

Multiple linear regression (MLR) is a research design that addresses these issues directly – which tells us how much of the variance in a dependent variable is accounted for by any group of independent variables, and which also tells us how important each independent variable is in accounting for this variance. In this chapter, we now look at MLR as a general strategy and design. In Chapter 12, we look at MLR as a general data analysis strategy.

Multiple linear regression (MLR) as a general strategy and design

Multiple linear regression – often abbreviated to MLR or just regression analysis – is basically a statistical technique for the analysis of data, but here I want to consider it as a strategy and design, as a way of conceptualising and organising quantitative research. It fits situations where we want to focus on a dependent variable, and to study its relationship with a number of independent variables. MLR is important because we want to do this sort of investigation very often. The conceptual framework is shown in Figure 10.3, and covariates may or may not be included. Of course, the conceptual framework is not limited to four independent variables. The general objective in the research is to account for variance in the dependent variable, and to see how the different independent variables, separately or in combination, contribute to accounting for this variance.

With MLR we can:

Estimate how much of the variance in a dependent variable we can account for using a particular set of independent variables. When most of the variance is accounted for, we are well on the way to understanding the dependent variable. Conversely, if only a small proportion of the variance is accounted for, we still have a long way to go in understanding it.

Determine the effects of the different independent variables on the dependent variable, by estimating the unique variance each independent variable accounts for.³ We can say which independent variables are of most and least importance in accounting for variance in the dependent variable, and therefore in bringing about change in the dependent variable. This is important knowledge when it comes to recommending strategies for changing the dependent variable.

Many quantitative research problems fit into this design, and many other studies can be designed in this way. Figure 10.4 shows a well-used example of this in

Independent variables

Dependent variable

Covariates

Y

FIGURE 10.3 Conceptual framework for multiple linear regression

education research (with four independent variables and no covariates), and Figure 10.5 shows several multiple-independent-variable-one-dependent-variable designs from different social science areas. Whenever a researcher is interested in relationships between variables, a regression analysis design can be used. There are also benefits to thinking this way about an overall research area. When (as often) the focus is on some major dependent variable, MLR provides a coordinated overall approach to the topic, as well as a ready-made conceptual framework, strategy and design. The requirements are that the researcher must be able to specify, define and measure the independent variables, in addition of course to the dependent variable.

Sometimes, the focus in research may be more on a detailed study of the relationship between variables, than on accounting for variance. That is a strategic decision to be made in planning the research, but it is also a matter of emphasis in how we think about the question, since these are two sides of the same coin. We account for variance by studying relationships with other variables. Therefore, even when the focus is on relationships, we do well to use a regression analysis design, for two reasons – all aspects of the relationship can be studied within that design, as noted in Chapter 12; and knowing how much variance we can account for gives us a strong indication of how important the relationship is.

To sum up, the conceptual framework that goes with MLR is useful because it addresses directly questions of key substantive significance. It deals with central questions of social science research interest, those that derive directly from causation. It also has two other advantages. First, it is flexible, in being able to accommodate different conceptual arrangements among the independent variables, including their joint effects on a dependent variable. This applies particularly to covariance analysis, interaction effects and non-linearity. As Chapter 12 will show, these are three important areas of research interest. Second, it is not a difficult approach to understand, conceptually or operationally. In this chapter we have been stressing its relevance to designing research, and noting that it comes with a ready-made set of research questions and conceptual framework. That is why it is described here as a general strategy and design. In Chapter 12, we look at MLR as a general data analysis strategy.

Independent variables

Intelligence

Socio-economic status

Motivation

Amount of homework

Dependent variable

School achievement

FIGURE

A regression analysis conceptual framework in education research

Examples of multiple-independent-variable

dependent-variable designs

Controlling variables

The term 'control' has already come up a number of times and it is another central concept in quantitative research design. It is extraneous variables we want to control – variables that may either confound the relationship

QUANTITATIVE RESEARCH DESIGN

In the design	In the analysis
Randomisation	Stratification
Restriction	Partial correlation
Matching	Analysis of covariance

we want to study or cause spurious interpretations of this relationship. To control such a variable means to remove its effects, or not to let it have any influence. There are two general ways variables are controlled in research – physical ways and statistical ways. In physical control, variables are controlled in the design. In statistical control, variables are controlled in the analysis of data. We will discuss each, in turn, and they are summarised in Table 10.1. Physical control is characteristic of experimental designs, whereas statistical control is more characteristic of correlational survey designs.

Physical control means that the variable is actually physically controlled, in the design of the study. There are three types of physical control:

Randomisation, where a variable can be controlled by having it vary randomly, or non-systematically. The Logic here is that the variable can have no systematic effect if it does not vary in a systematic way. Its effects will cancel each other out, because its variance is random. This idea is used when people are randomly assigned to treatment groups in a true experiment. As pointed out, this does not guarantee that the treatment groups will not differ from each other. Rather, it maximises the probability that they will not differ in any systematic way. *Restriction*, where a variable is controlled by physically restricting its variance and holding it constant in a study. Holding it constant means it has no variance in this piece of research. If it has no variance, it cannot have any covariance with other variables. This is the same as saying that it cannot show any relationship with other variables in the study, and therefore can have no effect on them. This second form of physical control is done in sample design and selection. For example, if gender was thought to be a possible extraneous or confounding factor in a study, one strategy would be to include in the study only either males or females. Because there is no variance in gender, it is controlled. The gain is clear – if only one sex group is included, then gender cannot be a factor in the relationships between other variables. But the Loss is clear too – if only males are included, the study can have nothing to say about females (or vice versa). This sort of trade-off occurs frequently in research. In this case, a more complete answer might be possible. Both sexes could be included, and then the sample could be partitioned into sex groups during the analysis of the data. This would achieve both control of gender and generalisability to both sex groups. It also shifts control of the variable from physical to statistical.

Matching, where group members are matched, one by one, on relevant characteristics. The problem, as noted, is that this quickly becomes impractical as the number of characteristics increases.

Statistical control means that the variable is not controlled in the design, but rather in the analysis of the data. Statistical control has to be designed into the study, in the sense that the variable to be controlled has to be measured. The logic of statistical control is that the analysis achieves a statistical approximation to the desired

(but unattainable) physical situation where the comparison groups are the same on the control variable. There are three types of statistical control:

stratification, or test factor elaboration (Rosenberg, 1968), where the control variable, the test factor, is partitioned, and the analysis of the relationship between variables is conducted within each level of the control variable, as described in the case of gender above.

partial correlation, where the variable to be controlled is partialled out of the relationship between the other two variables; this is a suitable technique when all variables are continuous, and is discussed in Chapter 11.

The analysis of covariance, where the control variable (the covariate) is extracted or co-varied out first, before the mainstream analysis is done.

The third of these ways, covariance analysis, has been described already, and is covered in more detail in Chapter 12, where it is incorporated into the MLR approach to data analysis. This third way is stressed here because it reflects the situation described in Sections 7 and 8 of this chapter – the very common situation in research, where there are several independent variables, one or more control variables and a dependent variable. While analysis of covariance is stressed, these three methods are logically equivalent to each other, in the sense that they are all designed to achieve the same thing – controlling the effects of unwanted or extraneous variables.

Chapter summary

Research design sits between research questions and data, showing how the research questions will be connected to the data, and what tools and procedures to use in answering them; design is based on strategy.

Quantitative research is fundamentally concerned with the relationship between variables; this is done either by comparing groups using experimental and quasi-experimental designs or by using non-experimental reasoning in correlational survey designs.

Independent variables (treatment variables, or 'causes') are manipulated in experiments to study their effects on dependent variables (outcome variables, or 'effects'); more generally, an independent variable is seen as the cause, and the dependent variable as the effect, in any cause-effect relationship.

Control variables, or covariates, are extraneous variables whose effects we want to remove or control, in order to see independent-dependent variable relationships more clearly.

A true experiment involves the manipulation of one or more independent variables and the random assignment of participants to comparison groups.

Quasi-experiments take advantage of naturally occurring treatment groups to study independent-dependent variable relationships, using the logic of experimental design.

Non-experimental designs – correlational surveys – take advantage of naturally occurring variation in independent variables, to study their relationship with dependent variables.

Variance is a central concept in quantitative research, and accounting for variance in a dependent variable is an important strategy for investigating causation.

Multiple linear regression directly addresses the question of how much dependent variable variance is accounted for by a set of independent variables.

Controlling variables is either done physically, through the research design, or statistically, through some form of covariance analysis.

KEY TERMS

Independent variable: the variable seen as the cause in a cause-effect relationship

Dependent variable: the variable seen as the effect in a cause-effect relationship

Control variable: an extraneous variable whose effects we want to remove or control; also called a covariate

Experiment: a technical research design term where one or more independent variables are manipulated to study their effect on a dependent variable, and where participants are randomly assigned to treatment or comparison groups

Quasi-experiment: a design which uses naturally occurring treatment groups to study independent-dependent variable relationships; uses the logic of experimental design

Correlational survey: a design which uses naturally occurring variation in independent variables to study relationships with dependent variables

Accounting for variance: a central strategy for quantitative research, which aims to account for the variation in a dependent variable through its relationships with independent variables

Multiple linear regression: a quantitative design with several independent variables and one dependent variable; estimates how much variance in the dependent variable is accounted for by these independent variables

Exercises and study questions

Define independent variable, dependent variable, control variable, and give examples of each. Sketch the design of an education experiment to compare students' learning under a new method of teaching (used with an experimental group) with students' learning under an old method of teaching (used with a control group). What design issues arise? Which variables might need to be controlled?

What is a quasi-experiment, and what is meant by naturally occurring treatment groups? Illustrate by comparing (a) educational achievement in big classes versus small classes, and (b) self-concept in children from intact families versus broken families.

What does accounting for variance in a dependent variable mean? What is its relationship to causation and why is it a central strategy in quantitative research? In conceptual terms, how is it done?

Draw a diagram to show the conceptual framework of a study with six independent variables and one dependent variable. What technique is appropriate for analysing the data from such a design?

6. What are the advantages of multiple linear regression as a general research design strategy?
7. What does controlling a variable mean? Why is it important in quantitative research?
8. Explain the logic of each type of control shown in Table 10.1.

Further reading

- Babbie, E. (2012) *The Practice of Social Research*. 13th edn. Belmont, CA: Wadsworth.
- Blalock, H.M. (1969) *Theory Construction: From Verbal to Mathematical Formulations*. Englewood Cliffs, NJ: Prentice-Hall.
- Brown, S.R. and Melamed, L. (1990) *Experimental Design and Analysis*. Newbury Park, CA: SAGE.
- Campbell, D.T. and Stanley, J.C. (1963) *Experimental and Quasi-Experimental Designs for Research*. Chicago, IL: Rand McNally.
- Cook, T.D. and Campbell, D.T. (1979) *Quasi-experimentation: Design and Analysis/Issues for Field Settings*. Chicago, IL: Rand McNally.
- Creswell, J.W. (2013) *Research Design: Qualitative and Quantitative Approaches*. 4th edn. Thousand Oaks, CA: SAGE.
- Devellis, O.A. (2013) *Surveys in Social Research*. 5th edn. London: Routledge.
- Fowler, F.J. (1988) *Survey Research Methods*. Newbury Park, CA: SAGE.
- Green, J.L., Camilli, G. and Elmore, P.B. (eds) (2006) *Handbook of Complementary Methods in Education Research*. Mahwah, NJ: Lawrence Erlbaum.
- Keppel, G. (1991) *Design and Analysis: A Researcher's Handbook*. 3rd edn. Englewood Cliffs, NJ: Prentice-Hall.
- Kerlinger, F.N. (1973) *Foundations of Behavioral Research*. New York: Holt, Rinehart and Winston.
- Marsh, C. (1982) *The Survey Method: The Contribution of Surveys to Sociological Explanation*. London: George Allen and Unwin.
- Sapsford, R. and Jupp, V. (1996) 'Validating evidence', in R. Sapsford and V. Jupp (eds). *Data Collection and Analysis*. London: SAGE. pp. 124.

Notes

This conceptual move, from discrete comparison groups to a continuum of variation, is actually an important and recurring theme in quantitative research. It comes up again in the discussion of measurement in Chapter 11.

- 2 Both simple descriptive surveys and correlational surveys are cross-sectional, with data collected from people at one point in time. Cross-sectional surveys need to be distinguished from longitudinal surveys, in which data are collected from people at different points over a period of time. Longitudinal research is an important specialised area – see Menard (1991).
- 1 In addition to estimating the unique contribution of variables, we can also estimate their joint contribution and any interaction effects.

May 1992

Quantitative Data Analysis: An Introduction

Preface

GAO assists congressional decisionmakers in their deliberative process by furnishing analytical information on issues and options under consideration. Many diverse methodologies are needed to develop sound and timely answers to the questions that are posed by the Congress. To provide GAO evaluators with basic information about the more commonly used methodologies, GAO's policy guidance includes documents such as methodology transfer papers and technical guidelines.

This methodology transfer paper on quantitative data analysis deals with information expressed as numbers, as opposed to words, and is about statistical analysis in particular because most numerical analyses by GAO are of that form. The intended reader is the GAO generalist, not statisticians and other experts on evaluation design and methodology. The paper aims to bridge the communications gap between generalist and specialist, helping the generalist evaluator be a wiser consumer of technical advice and helping report reviewers be more sensitive to the potential for methodological errors. The intent is thus to provide a brief tour of the statistical terrain by introducing concepts and issues important to GAO's work, illustrating the use of a variety of statistical methods, discussing factors that influence the choice of methods, and offering some advice on how to avoid pitfalls in the analysis of quantitative data. Concepts are presented in a nontechnical way by avoiding computational procedures, except for a few illustrations, and by avoiding a rigorous discussion of assumptions that underlie statistical methods.

Quantitative Data Analysis is one of a series of papers issued by the Program Evaluation and Methodology Division (PEMD). The purpose of the series is to provide GAO evaluators with guides to various

aspects of audit and evaluation methodology, to illustrate applications, and to indicate where more detailed information is available.

We look forward to receiving comments from the readers of this paper. They should be addressed to Eleanor Chelimsky at 202-275-1854.

A handwritten signature in black ink, appearing to read 'Werner Grosshans', with a large, sweeping loop at the end.

Werner Grosshans
Assistant Comptroller General
Office of Policy

A handwritten signature in black ink, appearing to read 'Eleanor Chelimsky', with a large, sweeping loop at the end.

Eleanor Chelimsky
Assistant Comptroller General
for Program Evaluation and Methodology

Contents

Preface		1
Chapter 1		8
Introduction	Guiding Principles	8
	Quantitative Questions Addressed in the Chapters of This Paper	11
	Attributes, Variables, and Cases	13
	Level of Measurement	16
	Unit of Analysis	18
	Distribution of a Variable	19
	Populations, Probability Samples, and Batches	26
	Completeness of the Data	28
	Statistics	29
Chapter 2		31
Determining the Central Tendency of a Distribution	Measures of the Central Tendency of a Distribution	33
	Analyzing and Reporting Central Tendency	35
Chapter 3		39
Determining the Spread of a Distribution	Measures of the Spread of a Distribution	41
	Analyzing and Reporting Spread	49
Chapter 4		51
Determining Association Among Variables	What Is an Association Among Variables?	51
	Measures of Association Between Two Variables	55
	The Comparison of Groups	67
	Analyzing and Reporting the Association Between Variables	70

Contents

Chapter 5		74
Estimating	Histograms and Probability Distributions	76
Population	Sampling Distributions	80
Parameters	Population Parameters	83
	Point Estimates of Population Parameters	84
	Interval Estimates of Population Parameters	87
Chapter 6		91
Determining	What Do We Mean by Causal Association?	92
Causation	Evidence for Causation	93
	Limitations of Causal Analysis	103
Chapter 7		105
Avoiding Pitfalls	In the Early Planning Stages	105
	When Plans Are Being Made for Data Collection	108
	As the Data Analysis Begins	109
	As the Results Are Produced and Interpreted	112
Appendixes	Bibliography	114
	Glossary	120
	Contributors	129
	Papers in This Series	130
Tables	Table 1.3: Generic Types of Quantitative Questions	11
	Table 1.1: Data Sheet for a Study of College Student Loan Balances	15
	Table 1.2: Tabular Display of a Distribution	26
	Table 2.1: Distribution of Staff Turnover Rates in Long-Term Care Facilities	32
	Table 2.2: Three Common Measures of Central Tendency	33
	Table 2.3: Illustrative Measures of Central Tendency	36
	Table 3.1: Measures of Spread	41
	Table 4.1: Data Sheet With Two Variables	52

Contents

	Table 4.2: Cross-Tabulation of Two Ordinal Variables	53
	Table 4.3: Percentaged Cross-Tabulation of Two Ordinal Variables	54
	Table 4.4: Cross-Tabulation of Two Nominal Variables	57
	Table 4.5: Two Ordinal Variables Showing No Association	70
	Table 5.1: Data Sheet for 100 Samples of College Students	81
	Table 5.2: Point and Interval Estimates for a Set of Samples	88

Figures	Figure 1.1: Histogram of Loan Balances	20
	Figure 1.2: Two Distributions	22
	Figure 1.3: Histogram for a Nominal Variable	25
	Figure 3.1: Histogram of Hospital Mortality Rates	40
	Figure 3.2: Spread of a Distribution	44
	Figure 3.3: Spread in a Normal Distribution	48
	Scatter Plots for Spending Level and Test Scores	59
	Regression of Test Scores on Spending Level	63
	Figure 4.3: Regression of Spending Level on Test Scores	65
	Figure 4.4: Linear and Nonlinear Associations	72
	Figure 5.1: Frequency Distribution of Loan Balances	76
	Figure 5.2: Probability Distribution of Loan Balances	78
	Figure 5.3: Sampling Distribution for Mean Student Loan Balances	82
	Figure 6.1: Causal Network	96

Contents

Abbreviations

AIDS	Acquired immune deficiency syndrome
GAO	U.S. General Accounting Office
PEMD	Program Evaluation and Methodology Division
PRE	Proportionate reduction in error
WIC	Special Supplemental Food Program for Women, Infants, and Children

Introduction

Guiding Principles

Data analysis is more than number crunching. It is an activity that permeates all stages of a study. Concern with analysis should (1) begin during the design of a study, (2) continue as detailed plans are made to collect data in different forms, (3) become the focus of attention after data are collected, and (4) be completed only during the report writing and reviewing stages.¹

The basic thesis of this paper is that successful data analysis, whether quantitative or qualitative, requires (1) understanding a variety of data analysis methods, (2) planning data analysis early in a project and making revisions in the plan as the work develops; (3) understanding which methods will best answer the study questions posed, given the data that have been collected; and (4) once the analysis is finished, recognizing how weaknesses in the data or the analysis affect the conclusions that can properly be drawn. The study questions govern the overall analysis, of course. But the form and quality of the data determine what analyses can be performed and what can be inferred from them. This implies that the evaluator should think about data analysis at four junctures:

- when the study is in the design phase,
- when detailed plans are being made for data collection,
- after the data are collected, and
- as the report is being written and reviewed.

Designing the Study

As policy-relevant questions are being formulated, evaluators should decide what data will be needed to

¹Relative to GAO job phases, the first two checkpoints occur during the job design phase, the third occurs during data collection and analysis, and the fourth during product preparation. For detail on job phases see the General Policy Manual, chapter 6, and the Project Manual, chapters 6.2, 6.3, and 6.4.

answer the questions and how they will analyze the data. In other words, they need to develop a data analysis plan. Determining the type and scope of data analysis is an integral part of an overall design for the study. (See the transfer paper entitled Designing Evaluations, listed in “Papers in This Series.”) Moreover, confronting data collection and analysis issues at this stage may lead to a reformulation of the questions to ones that can be answered within the time and resources available.

Data Collection

When evaluators have advanced to the point of planning the details of data collection, analysis must be considered again. Observations can be made and, if they are qualitative (that is, text data), converted to numbers in a variety of ways that affect the kinds of analyses that can be performed and the interpretations that can be made of the results. Therefore, decisions about how to collect data should be influenced by the analysis options in mind.

Data Analysis

After the data are collected, evaluators need to see whether their expectations regarding data characteristics and quality have been met. Choice among possible analyses should be based partly on the nature of the data—for example, whether many observed values are small and a few are large and whether the data are complete. If the data do not fit the assumptions of the methods they had planned to use, the evaluators have to regroup and decide what to do with the data they have.² A different form of data analysis may be advisable, but if some

²An example would be a study in which the data analysis method evaluators planned to use required the assumption that observations be from a probability sample, as discussed in chapter 5. If the evaluators did not obtain observations for a portion of the intended sample, the assumption might not be warranted and their application of the method could be questioned.

observations are untrustworthy or missing altogether, additional data collection may be necessary.

As the evaluators proceed with data analysis, intermediate results should be monitored to avoid pitfalls that may invalidate the conclusions. This is not just verifying the completeness of the data and the accuracy of the calculations but maintaining the logic of the analysis. Yet it is more, because the avoidance of pitfalls is both a science and an art. Balancing the analytic alternatives calls for the exercise of considerable judgment. For example, when observations take on an unusual range of values, what methods should be used to describe the results? What if there are a few very large or small values in a set of data? Should we drop data at the extreme high and low ends of the scale? On what grounds?

**Writing and
Reviewing**

Finally, as the evaluators interpret the results and write the report, they have to close the loop by making judgments about how well they have answered the questions, determining whether different or supplementary analyses are warranted, and deciding the form of any recommendations that may be suitable. They have to ask themselves questions about their data collection and analysis: How much of the variation in the data has been accounted for? Is the method of analysis sensitive enough to detect the effects of a program? Are the data “strong” enough to warrant a far-reaching recommendation? These questions and many others may occur to the evaluators and reviewers and good answers will come only if the analyst is “close” to the data but always with an eye on the overall study questions.

Quantitative Questions Addressed in the Chapters of This Paper

Most GAO statistical analyses address one or more of the four generic questions presented in table 1.3. Each generic question is illustrated with several specific questions and examples of the kinds of statistics that might be computed to answer the questions. The specific questions are loosely based on past GAO studies of state bottle bills (U.S. General Accounting Office, 1977 and 1980).

Table 1.3: Generic Types of Quantitative Questions

Generic question	Specific question	Useful statistics
What is a typical value of the variable?	At the state level, how many pounds of soft drink bottles (per unit of population) were typically returned annually?	Measures of central tendency (ch. 2)
How much spread is there among the cases? To what extent are two or more variables associated?	How similar are the individual states' return rates? What factors are most associated with high return rates: existence of state bottle bills? state economic conditions? state levels of environmental awareness?	Measures of spread (ch. 3) Measures of association (ch. 4)
To what extent are there causal relationships among two or more variables?	What factors cause high return rates: existence of state bottle bills? state economic conditions? state level of environmental awareness?	Measures of association (ch. 4): Note that association is but one of three conditions necessary to establish causation (ch. 6)

Bottle bills have been adopted by about nine states and are intended to reduce solid waste disposal problems by recycling. Other benefits can also be sought, such as the reduction of environmental litter and savings of energy and natural resources. One of GAO's studies was a prospective analysis, intended to inform discussion of a proposed national bottle bill. The quantitative analyses were not the only relevant

factor. For example, the evaluators had to consider the interaction of the merchant-based bottle bill strategy with emerging state incentives for curbside pickups or with other recycling initiatives sponsored by local communities. The quantitative results were, however, relevant to the overall conclusions regarding the likely benefits of the proposed national bottle bill.

The first three generic questions in table 1.3 are standard fare for statistical analysis. GAO reports using quantitative analysis usually include answers in the form of descriptive statistics such as the mean, a measure of central tendency, and the standard deviation, a measure of spread. In chapters 2, 3, and 4 of this paper, we focus on descriptive statistics for answering the questions.

To answer many questions, it is desirable to use probability samples to draw conclusions about populations. In chapter 5, we address the first three questions from the perspective of inferential statistics. The treatment there is necessarily brief, focused on point and interval estimation methods.

The fourth generic question, about causality, is more difficult to answer than the others. Providing a good answer to a causal question depends heavily upon the study design and somewhat advanced statistical methods; we treat the topic only lightly in chapter 6. Chapter 7 discusses some broad strategies for avoiding pitfalls in the analysis of quantitative data.

Before describing these concepts, it is important to establish a common understanding about some ideas that are basic to data analysis, especially those applicable to the quantitative analysis we describe in this paper. Each of GAO's assignments requires considerable analysis of data. Over the years, many

workable tools and methods have been developed and perfected. Trained evaluators use these tools as appropriate in addressing an assignment's objectives. This paper tries to reinforce the uses of these tools and put consistent labels on them.³ It also gives helpful hints and illustrates the use of each tool. In the next section, we discuss the basic terminology that is used in later chapters.

Attributes, Variables, and Cases

Observations about persons, things, and events are central to answering questions about government programs and policies. Groups of observations are called data, which may be qualitative or quantitative. Statistical analysis is the manipulation, summarization, and interpretation of quantitative data.

We observe characteristics of the entities we are studying. For example, we observe that a person is female and we refer to that characteristic as an attribute of the person. A logical collection of attributes is called a variable; in this instance, the variable would be gender and would be composed of the attributes female and male.⁴ Age might be another variable composed of the integer values from 0 to 115.

³Inconsistencies in the use of statistical terms can cause problems. We have tried to deal with the difficulty in three ways: (1) by using the language of current writers in the field, (2) by noting instances where there are common alternatives to key terms, and (3) by including a glossary of the terms used in this paper.

⁴Instead of referring to the attributes of a variable, some prefer to say that the variable takes on a number of "values." For example, the variable gender can have two values, male and female. Also, some statisticians use the expression "attribute sampling" in reference to probability sampling procedures for estimating proportions. Although attribute sampling is related to attribute as used in data analysis, the terminology is not perfectly parallel. See the discussion of attribute sampling in the transfer paper entitled Using Statistical Sampling, listed in "Papers in This Series."

It is convenient to refer to the variables we are especially interested in as response variables. For example, in a study of the effects of a government retraining program for displaced workers, employment rate might be the response variable. In trying to determine the need for an acquired immune deficiency syndrome (AIDS) education program in different segments of the U.S. population, evaluators might use the incidence of AIDS as the response variable. We usually also collect information on other variables with which we hope to better understand the response variables. We occasionally refer to these other variables as supplementary variables.

The data that we want to analyze can be displayed in a rectangular or matrix form, often called a data sheet (see table 1.1). To simplify matters, the individual persons, things, or events that we get information about are referred to generically as cases. (The intensive study of one or a few cases, typically combining quantitative and qualitative data, is referred to as case study research. See the GAO transfer paper entitled Case Study Evaluations.) Traditionally, the rows in a data sheet correspond to the cases and the columns correspond to the variables of interest. The numbers or words in the cells then correspond to the attributes of the cases.

**Table 1.1: Data Sheet for
a Study of College
Student Loan Balances**

Case	Age	Class	Type of institution	Loan balance
1	23	Sophomore	Private	\$3,254
2	19	Freshman	Public	1,501
3	21	Junior	Public	2,361
4	30	Graduate	Private	8,100
5	21	Freshman	Private	1,970
6	22	Sophomore	Public	3,866
7	21	Sophomore	Public	2,890
8	20	Freshman	Public	6,300
9	22	Junior	Private	2,639
10	21	Sophomore	Public	1,718
11	19	Freshman	Private	2,690
12	20	Sophomore	Public	3,812
13	20	Sophomore	Public	2,210
14	23	Senior	Private	3,780
15	24	Senior	Private	5,082

Table 1.1 shows 15 cases, college students, from a hypothetical study of student loan balances at higher education institutions. The first column shows an identification number for each case, and the rest of the columns indicate four variables: age of student, class, type of institution, and loan balance. Two of the variables, class and type of institution, are presently in text form. As will be seen shortly, they can be converted to numbers for purposes of quantitative analysis. Loan balance is the response variable and the others are supplementary.

The choice of a data analysis method is affected by several considerations, especially the level of measurement for the variables to be studied; the unit of analysis; the shape of the distribution of a variable,

including the presence of outliers (extreme values); the study design used to produce the data from populations, probability samples, or batches; and the completeness of the data. Each factor is considered briefly.

Level of Measurement

Quantitative variables take several forms, frequently called levels of measurement, which affect the type of data analysis that is appropriate. Although the terminology used by different analysts is not uniform, one common way to classify a quantitative variable is according to whether it is nominal, ordinal, interval, or ratio.

The attributes of a nominal variable have no inherent order. For example, gender is a nominal variable in that being male is neither better nor worse than being female. Persons, things, and events characterized by a nominal variable are not ranked or ordered by the variable. For purposes of data analysis, we can assign numbers to the attributes of a nominal variable but must remember that the numbers are just labels and must not be interpreted as conveying the order of the attributes. In the study of student loans, the type of institution is a nominal variable with two attributes—private and public—to which we might assign the numbers 0 and 1 or, if we wish, 12 and 17. For most purposes, 0 and 1 would be more useful.⁵

With an ordinal variable, the attributes are ordered. For example, observations about attitudes are often arrayed into five classifications, such as greatly dislike, moderately dislike, indifferent to, moderately like, greatly like. Participants in a government program might be asked to categorize their views of the program offerings in this way. Although the

⁵A variable for which the attributes are assigned arbitrary numerical values is usually called a “dummy variable.” Dummy variables occur frequently in evaluation studies.

ordinal level of measurement yields a ranking of attributes, no assumptions are made about the “distance” between the classifications. In this example, we do not assume that the difference between persons who greatly like a program offering and ones who moderately like it is the same as the difference between persons who moderately like the offering and ones who are indifferent to it. For data analysis, numbers are assigned to the attributes (for example, greatly dislike = -2, moderately dislike = -1, indifferent to = 0, moderately like = +1, and greatly like = +2), but the numbers are understood to indicate rank order and the “distance” between the numbers has no meaning. Any other assignment of numbers that preserves the rank order of the attributes would serve as well. In the student loan study, class is an ordinal variable.

The attributes of an interval variable are assumed to be equally spaced. For example, temperature on the Fahrenheit scale is an interval variable. The difference between a temperature of 45 degrees and 46 degrees is taken to be the same as the difference between 90 degrees and 91 degrees. However, it is not assumed that a 90-degree object has twice the temperature of a 45-degree object (meaning that the ratio of temperatures is not necessarily 2 to 1). The condition that makes the ratio of two observations uninterpretable is the absence of a true zero for the variable. In general, with variables measured at the interval level, it makes no sense to try to interpret the ratio of two observations.

The attributes of a ratio variable are assumed to have equal intervals and a true zero point. For example, age is a ratio variable because the negative age of a person or object is not meaningful and, thus, the birth of the person or the creation of the object is a true zero point. With ratio variables, it makes sense to

form ratios of observations and it is thus meaningful, for example, to say that a person of 90 years is twice as old as one of 45. In the study of student loans, age and loan balance are both ratio variables (the attributes are equally spaced and the variables have true zero points). For analysis purposes, it is seldom necessary to distinguish between interval and ratio variables so we usually lump them together and call them interval-ratio variables.

Unit of Analysis

Units of analysis are the persons, things, or events under study—the entities that we want to say something about. Frequently, the appropriate units of analysis are easy to select. They follow from the purpose of the study. For example, if we want to know how people feel about the offerings of a government program, individual people would be the logical unit of analysis. In the statistical analysis, the set of data to be manipulated would be variables defined at the level of the individual.

However, in some studies, variables can potentially be analyzed at two or more levels of aggregation. Suppose, for example, that evaluators wished to evaluate a compensatory reading program and had acquired reading test scores on a large number of children, some who participated in the program and some who did not. One way to analyze the data would be to treat each child as a case.

But another possibility would be to aggregate the scores of the individual children to the classroom level. For example, they could compute the average scores for the children in each classroom that participated in their study. They could then treat each classroom as a unit, and an average reading test score would be an attribute of a classroom. Other variables, such as teacher's years of experience, number of

students, and hours of instruction could be defined at the classroom level. The data analysis would proceed by using classrooms as the unit of analysis. For some issues, treating each child as a unit might seem more appropriate, while in others each classroom might seem a better choice. And we can imagine rationales for aggregating to the school, school district, and even state level.

Summarizing, the unit of analysis is the level at which analysis is conducted. We have, in this example, five possible units of analysis: child, classroom, school, school district, and state. We can move up the ladder of aggregation by computing average reading scores across lower-level units. In effect, the definition of the variable changes as we change the unit of analysis. The lowest-level variable might be called child-reading-score, the next could be classroom-average-reading-score, and so on.

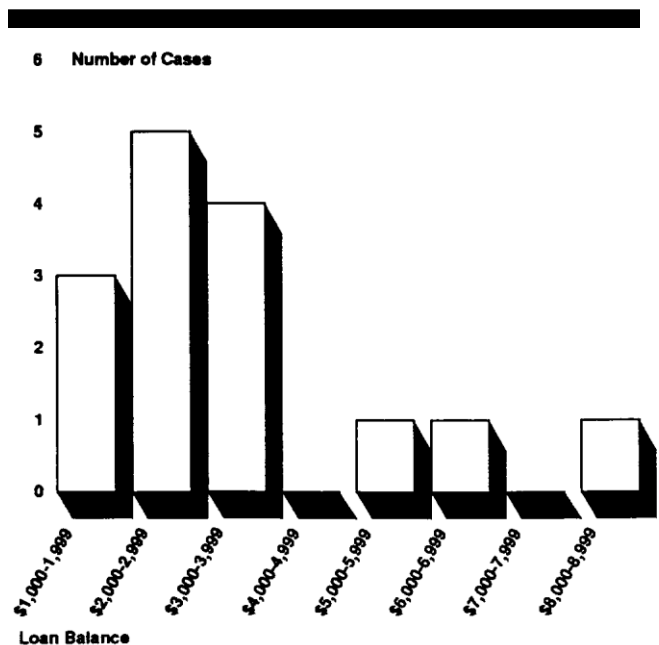
In general, the results from an analysis will vary, depending upon the unit of analysis. Thus, for studies in which aggregation is a possibility, evaluators must answer the question: What is the appropriate unit of analysis? Several situation-specific factors may need consideration, and there may not be a clear-cut answer. Sometimes analyses are carried out with several units of analysis. (GAO evaluators should seek advice from technical assistance groups.)

Distribution of a Variable

The cases we observe vary in the characteristics of interest to us. For example, students vary by class and by loan balance. Such variation across cases, which is called the distribution of a variable, is the focus of attention in a statistical analysis. Among the several ways to picture or describe a distribution, the histogram is probably the simplest. To illustrate, suppose we want to display the distribution of the

loan balance variable for the 15 cases in table 1.1. A histogram for the data is shown in figure 1.1. The length of the left hand bar corresponds to the number of observations between \$1,000 and \$1,999. There are three: \$1,500, \$1,970, and \$1,718. The lengths of the other bars are determined in a similar fashion, and the overall histogram gives a picture of the distribution. In this example, the distribution is rather “piled up” on one end and spread out at the other; two intervals have no observations.

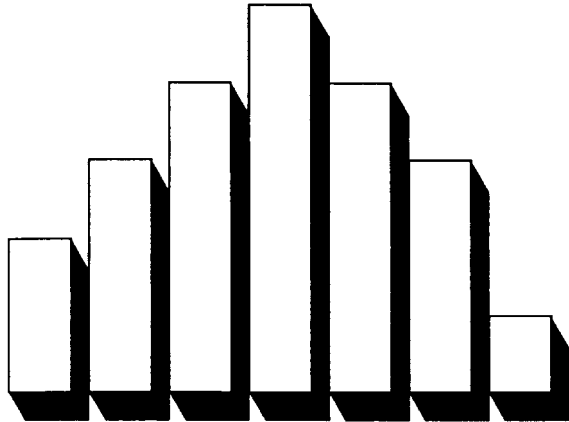
Figure 1.1: Histogram of Loan Balances



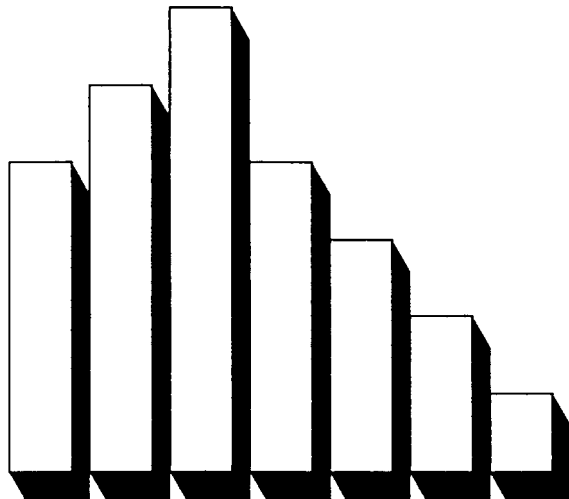
Histograms show the shape of a distribution, a factor that helps determine the type of data analysis that will

be appropriate. For example, some techniques are suitable only when the distribution is approximately symmetrical (as in figure 1.2a), while others can be

**Figure 1.2: Two
Distributions**



a. Approximately Symmetrical Distribution



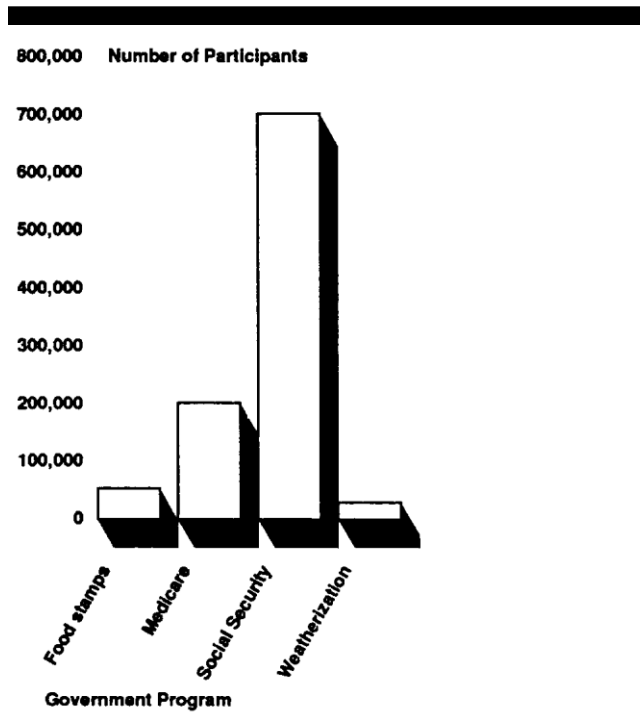
b. Asymmetrical Distribution

used when the observations are asymmetrical (figure 1.2b). Once data are collected for a study, we need to inspect the distributions of the variables to see what initial steps are appropriate for the data analysis. Sometimes it is advisable to transform a variable (that is, systematically change the values of the observations) that is distributed asymmetrically to one that is symmetric. For example, taking the square root of each observation is a transformation that will sometimes work. Velleman and Hoaglin (1981, ch. 2) provide a good introduction to transformation strategies (they refer to them as “re-expression”) and Hoaglin, Mosteller, and Tukey (1983, ch. 4) give a more complete treatment. (GAO generalists who believe that such a strategy is in order are advised to seek help from a technical assistance group.) With proper care, transformations do not alter the conclusions that can be drawn from data.

Another aspect of a distribution is the possible presence of outliers, a few observations that have extremely large or small values so that they lie on the outer reaches of the distribution. For the student loan observations, case number 4, which has a value of \$8,100, is far from the center of the distribution. Outliers can be important because they may lead to new understanding of the variable in question. However, outliers attributable to measurement error may produce misleading results with some statistical analyses, so an early decision must be made about how to handle outliers—a decision not easy to make. The usual way is to employ analytical methods that are relatively insensitive to outliers—for example, by using the median instead of the mean. Sometimes outliers are dropped from the analysis but only if there is good reason to believe that the observations are in error.

Considerations about the shape of a distribution and about outliers apply to ordinal, interval, and ratio variables. Because the attributes of a nominal variable have no inherent order, these spatial relationships have no meaning. However, we can still display the results from observations on a nominal variable as a histogram, as long as we remember that the order of the attributes is arbitrary. Figure 1.3 shows hypothetical data on the number of participants in four government programs. There is no inherent order for displaying the programs.

Figure 1.3: Histogram for
a Nominal Variable



Another way of showing the distribution of a variable is to use a simple table. Suppose evaluators have data on 341 homeowners' attitudes toward energy conservation with three categories of response: indifferent, somewhat positive, and positive. Table 1.2 shows the data in summary form. This kind of display is not often used when only one variable is involved, but with two it is common (see chapter 4).

Table 1.2: Tabular
Display of a Distribution

Attitude toward energy conservation	Number of homeowners
Indifferent	120
Somewhat positive	115
Positive	106
Total	341

Populations,
Probability
Samples, and
Batches

Statistical analysis is applied to a group of cases. The process by which the group was chosen (that is, the study design) affects the type of data analysis that is appropriate and the interpretations that may be drawn from the analysis. Three types of group are of interest: populations, probability samples, and batches.

A population is the full set of cases that the evaluators have a question about. For example, suppose they want to know the age of Medicaid participants and the amount of benefits these participants received last year. The population would be all persons who received such benefits, and the evaluators might obtain data tapes containing the attributes for all such persons. They could perform statistical analyses to describe the distributions of certain variables such as age and amount of benefits received. The results of such an analysis are called descriptive statistics.

A second way to draw conclusions about the Medicaid participants is to use a probability sample from the population of beneficiaries. A probability sample is a group of cases selected so that each member of the population has a known, nonzero probability of being selected. (For detailed information on probability sampling, see the transfer paper entitled Using Statistical Sampling.) Studies based on probability samples are usually less

expensive than those that use data from the entire population and, under some conditions, are less error-prone.⁶ The study of probability samples can use descriptive statistics but the study of the population, upon which the probability sample is based, uses inferential statistics (discussed in chapter 5).

A group of cases can also be treated as a batch, a group produced by a process about which we make no probabilistic assumptions. For example, the evaluators might use their judgment, not probability, to select a number of interesting Medicaid cases for study. Being neither a population nor a probability sample, the set of cases is treated as a batch. As such, the techniques of descriptive statistics can be applied but not those of inferential statistics. Thus, conclusions about the population of which the batch is a part cannot be based on statistical rules of inference.

When do we regard a group of cases as a batch? Evaluators who have purposely chosen a nonprobability sample, or who have doubts about whether cases in hand fit the definition of a probability sample—for example, because they are using someone else’s data and the selection procedures were not well described—should treat the cases as a batch. Actually, any group of cases can be regarded as a batch. The term is applied whenever we do not wish to assume the grouping is a population or a probability sample.

⁶Error in using probability samples to answer questions about populations stems from the net effects of both measurement error and sampling error. Conclusions based upon data from the entire population are subject only to measurement error. The total error associated with data from a probability sample may be less than the total error (measurement only) of data from a population.

Completeness of the Data

When we design a study, we plan to obtain data for a specific number of cases. Despite our best plans, we usually cannot obtain data on all variables for all cases. For example, in a sample survey, some persons may decline to respond at all and others may not answer certain questionnaire items. Or responses to some interview questions may be inadvertently “lost” during data editing and processing. In another study, we may not be allowed to observe certain events. Almost inevitably, the data will be incomplete in several respects, and data analysis must contend with that eventuality.

Incompleteness in the data can affect analysis in a variety of ways. The classic example is when we draw a probability sample with the aim of using inferential statistics to answer questions about a population. To illustrate, suppose evaluators send a questionnaire to a sample of Medicaid beneficiaries but only 45 percent provide data. Without increasing the response rate or satisfying themselves that nonrespondents would have answered in ways similar to respondents (or that the differences would have been inconsequential), the evaluators would not be entitled to draw inferential conclusions about the population of Medicaid beneficiaries. If they knew the views of the nonrespondents, their overall description of the population might be quite different. They would be limited, therefore, to descriptive statistics about the 45 percent who responded, and that information might not be useful for answering a policy-relevant question.

The problem of incomplete data entails several considerations and a variety of analytic approaches. (See, for example, Groves, 1989; Madow, Olkin, and Rubin, 1983; and Little and Rubin, 1987.) One important strategy is to minimize the problems by using good data collection techniques. (See the

transfer papers entitled Using Structured Interviewing Techniques and Developing and Using Questionnaires.)

Statistics

In GAO work, we may be interested in analyzing data from a population, a probability sample, or a batch. Regardless of how the group of cases is selected, we make observations on the cases and can produce a data sheet like that of table 1.1. A main purpose of statistical analysis is to draw conclusions about the real world by computing useful statistics.⁷ A statistic is a number computed from a set of data. For example, the midpoint loan balance for the 15 students, \$2,890, is a statistic—the median loan balance for the batch in statistical terminology.

Many statistics are possible but only a relative few are useful in the sense of helping us understand the data and answer policy-relevant questions. Another possibly useful statistic from the batch of 15 is the range—the difference between the maximum loan balance and the minimum. The range, in this example, is $8,100 - 1,500 = 6,600$. In this instance, the “computation” of the statistic is merely a sorting through the attributes for the loan balance variable to find the largest and smallest values and then computing the difference between them. Many statistics can be imagined but most would not be useful in describing the batch. For example, the square root of the difference between the maximum loan balance and the mean loan balance is a statistic but not a useful one.

The methods of statistical analysis provide us with ways to compute and interpret useful statistics. Those

⁷Another purpose, though one that has received less attention in the statistical literature, is to devise useful ways to graphically depict the data. See, for example, Du Toit, Steyn, and Stumpf, 1986; and Tufte, 1983.

that are useful for describing a population or a batch are called descriptive statistics. They are used to describe a set of cases upon which observations were made. Methods that are useful for drawing inferences about a population from a probability sample are called inferential statistics. They are used to describe a population using merely information from observations on a probability sample of cases from the population. Thus, the same statistic can be descriptive or inferential or both, depending on its use.

Determining the Central Tendency of a Distribution

Descriptive analyses are the workhorses of GAO, carrying much of the message in many of our reports. There are three main forms of descriptive analysis: determining the central tendency in the distribution of a variable (discussed in this chapter), determining the spread of a distribution (chapter 3), and determining the association among variables (chapter 4).

The determination of central tendency answers the first of GAO's four basic questions, What is a typical value of the variable? All readers are familiar with the basic ideas. Sample questions might be

- How satisfied are Social Security beneficiaries with the agency's responsiveness?
- How much time is required to fill requests for fighter plane repair parts?
- What was the dollar value in agricultural subsidies received by wealthy farmers?
- What was the turnover rate among personnel in long-term care facilities?

The common theme of these questions is the need to express what is typical of a group of cases. For example, in the last question, the response variable is the turnover rate. Suppose evaluators have collected information on the turnover rates for 800 long-term care facilities. Assuming there is variation among the facilities, they would have a distribution for the turnover rate variable. There are two approaches for describing the central tendency of a distribution: (1) presenting the data on turnover rates in tables or figures and (2) finding a single number, a descriptive statistic, that best summarizes the distribution of turnover rates.

The first approach, shown in table 2.1, allows us to "see" the distribution. The trouble is that it may be

hard to grasp what the typical value is. However, evaluators should always take a graphic or tabular approach as a first step to help in deciding how to proceed on the second approach, choosing a single statistic to represent the batch. How a display of the distribution can help will be seen shortly.

Table 2.1: Distribution of Staff Turnover Rates in Long-Term Care Facilities

Turnover rates (percent new staff per year)	Frequency count (number of long-term care facilities)
0-0.9	155
1.0-1.9	100
2.0-2.9	125
3.0-3.9	150
4.0-4.9	100
5.0-5.9	75
6.0-6.9	50
7.0-7.9	25
8.0-8.9	15
9.0-9.9	5

The second approach, describing the typical value of a variable with a single number, offers several possibilities. But before considering them, a little discussion of terminology is necessary. A descriptive statistic is a number, computed from observations of a batch, that in some way describes the group of cases. The definition of a particular descriptive statistic is specific, sometimes given as a recipe for calculation. Measures of central tendency form a class of descriptive statistics each member of which characterizes, in some sense, the typical value of a variable—the central location of a distribution.¹ The

¹Measures of central tendency also go by other, equivalent names such as “center indicators” and “location indicators.”

definition of central tendency is necessarily somewhat vague because it embraces a variety of computational procedures that frequently produce different numerical values. Nonetheless, the purpose of each measure would be to compress information about a whole distribution of cases into a single number.

Measures of the Central Tendency of a Distribution

Three familiar and commonly used measures of central tendency are summarized in table 2.2. The mean, or arithmetic average, is calculated by summing the observations and dividing the sum by the number of observations. It is ordinarily used as a measure of central tendency only with interval-ratio level data. However, the mean may not be a good choice if several cases are outliers or if the distribution is notably asymmetric. The reason is that the mean is strongly influenced by the presence of a few extreme values, which may give a distorted view of central tendency. Despite such limitations, the mean has definite advantages in inferential statistics (see chapter 5).

Table 2.2: Three Common Measures of Central Tendency

Measurement level	Use of measure ^a		
	Mode	Median	Mean
Nominal	Yes	No	No
Ordinal	Yes	Yes	No ^b
Interval-ratio	Yes	Yes	Yes ^c

^a"Yes" means the indicator is suitable for the measurement level shown.

^bMay be OK in some circumstances. See chapter 7.

^cMay be misleading when the distribution is asymmetric or has a few outliers.

The median—calculated by determining the midpoint of rank-ordered cases—can be used with ordinal, interval, or ratio measurements and no assumptions need be made about the shape of the distribution.² The median has another attractive feature: it is a resistant measure. That means it is not much affected by changes in a few cases. Intuitively, this suggests that significant errors of observation in several cases will not greatly distort the results. Because it is a resistant measure, outliers have less influence on the median than on the mean. For example, notice that the observations 1, 4, 4, 5, 7, 7, 8, 8, 9 have the same median (7) as the observations 1, 4, 4, 5, 7, 7, 8, 8, 542. The means (5.89 and 65.44, respectively), however, are quite different because of the outlier, 542, in the second set of observations.

The mode is determined by finding the attribute that is most often observed.³ That is, we simply count the number of times each attribute occurs in the data, and the mode is the most frequently occurring attribute. It can be used as a measure of central tendency with data at any level of measurement. However, the mode is most commonly employed with nominal variables and is generally less used for other levels. A distribution can have more than one mode (when two or more attributes tie for the highest frequency). When it does, that fact alone gives important information about the shape of the distribution.

Measures of central tendency are used frequently in GAO reports. In a study of tuition guarantee programs (U.S. General Accounting Office, 1990c), for example,

²With an odd number of cases, the midpoint is the median. With an even number of cases, the median is the mean of the middle pair of cases.

³This definition is suitable when the mode is used with nominal and ordinal variables—the most common situation. A slightly different definition is required for interval-ratio variables.

the mean was often used to characterize the programs in the sample, but when outliers were evident, the median was reported. In another GAO study (U.S. General Accounting Office, 1988), the distinctions between properties of the mode, median, and mean figured prominently in an analysis of procedures used by the Employment and Training Administration to determine prevailing wage rates of farmworkers.

Analyzing and Reporting Central Tendency

To illustrate some considerations involved in determining the central tendency of a distribution, we can recall the earlier study question about the views of Social Security beneficiaries regarding program services. Assume that a questionnaire has been sent to a batch of 800 Social Security recipients asking how satisfied they are with program services.⁴ Further, imagine four hypothetical distributions of the responses. By assigning a numerical value of 1 to the item response “very satisfied” and 5 to “very dissatisfied,” and so on, we can create an ordinal variable. The three measures of central tendency can then be computed to produce the results in table 2.3.⁵ Although the data are ordinal, we have included the mean for comparison purposes.

⁴To keep the discussion general, we make no assumptions about how the group of recipients was chosen. However, in GAO, a probability sample would usually form the basis for data collection by a mailout questionnaire.

⁵Although computer programs automatically compute a variety of indicators and although we display three of them here, we are not suggesting that this is a good practice. In general, the choice of an indicator should be based upon the measurement level of a variable and the shape of the distribution.

Table 2.3: Illustrative Measures of Central Tendency

Attribute	Code	Distribution			
		A	B	C	D
Very satisfied	1	250	250	100	159
Satisfied	2	200	150	150	159
Neither satisfied nor dissatisfied	3	125	0	300	164
Dissatisfied	4	125	150	150	159
Very dissatisfied	5	100	250	100	159
Total responses		800	800	800	800
Mean		2.5	3	3	3
Median		2	3	3	3
Mode		1 and 1 5 3 3			

In distribution A, the data are distributed asymmetrically. More persons report being very satisfied than any other condition, and mode 1 reflects this. However, 225 beneficiaries expressed some degree of dissatisfaction (codes 4 and 5), and these observations pull the mean to a value of 2.5, (that is, toward the dissatisfied end of the scale). The median is 2, between the mode and the mean. Although the mean might be acceptable for some ordinal variables, in this example it can be misleading and shows the danger of using a single measure with an asymmetrical distribution. The mode seems unsatisfactory also because, although it draws attention to the fact that more respondents reported satisfaction with the services than any other category, it obscures the point that 225 reported that they were dissatisfied or very dissatisfied. The median seems the better choice for this distribution if we can display only one number, but showing the whole distribution is probably wise.

In distribution B, the mean and the median both equal 3 (a central tendency of “neither satisfied nor dissatisfied”). Some would say this is nonsense in terms of the actual distribution, since no one actually chose the middle category. Modes 1 and 5 seem the better choices to represent the clearly bimodal distribution, although again a display of the full distribution is probably the best option.

In distribution C, the mean, median, and mode are identical; the distribution is symmetrical. Any one of the three would be appropriate. One easy check on the symmetry of a distribution, as this shows, is to compare the values of the mean, median, and mode. If they differ substantially, as with distribution A, the distribution is probably such that the median should be used.

As distribution D illustrates, however, this rule-of-thumb is not infallible. Although the mean, median, and mode agree, the distribution is almost flat. In this case, a single measure of central tendency could be misleading, since the values 1, 2, 3, 4, and 5 are all about equally likely to occur. Thus, the full distribution should be displayed.

The lesson of this example? First, before representing the central tendency by any single number, evaluators need to look at the distribution and decide whether the indicator would be misleading. Second, there will be occasions when displaying the results graphically or in tabular form will be desirable instead of, or in addition to, reporting statistics.

The interpretation of a measure of central tendency comes from the context of the associated policy question. The number itself does not carry along a message saying whether policymakers should be complacent or concerned about the central tendency.

Chapter 2
Determining the Central Tendency of a
Distribution

For example, the observed mean agricultural subsidy for farmers can be interpreted only in the context of economic and social policy. Comparison of the mean to other numbers such as the wealth or income level of farmers or to the trend over time for mean subsidies might be helpful in this regard. And, of course, limits on mean values are sometimes written into law. An example is the fleet-average mileage standard for automobiles. Information that can be used to interpret the observed measures of central tendency is a necessary part of the overall answer to a policy question.

Determining the Spread of a Distribution

Spread refers to the extent of variation among cases—sometimes cases cluster closely together and sometimes they are widely spread out. When we determine appropriate policy action, the spread of a distribution may be as much a factor, or more, than the central tendency.

The point is illustrated by the issue of variation in hospital mortality rates. Consider two questions. How much do hospital mortality rates vary? If there is substantial variation, what accounts for it? We consider questions of the first type in this chapter and questions of the second type in chapter 6.

Figure 3.1: Histogram of Hospital Mortality Rates

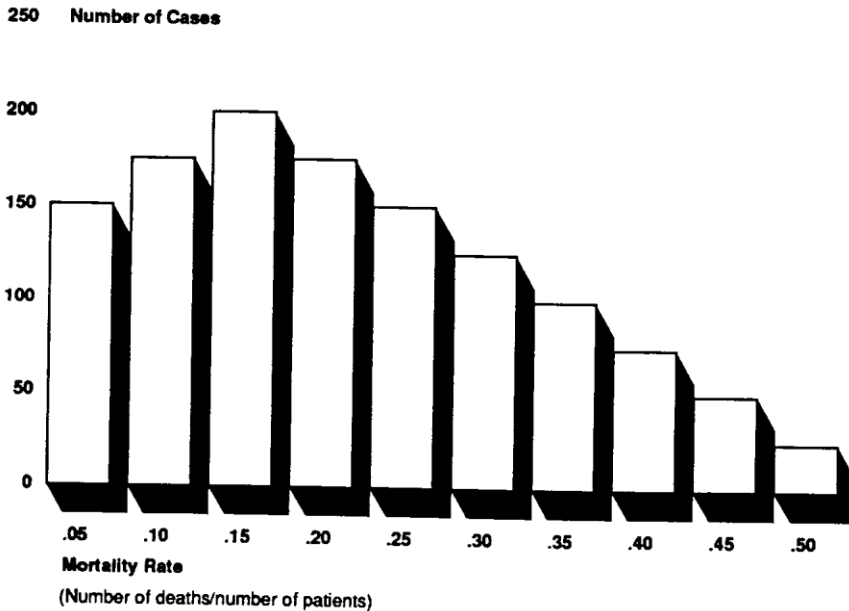


Figure 3.1 shows the distribution of hypothetical data on mortality rates in 1,225 hospitals. While the depiction is useful both in gaining an initial understanding of the spread in mortality rates and in communicating findings, it is also usually desirable to produce a number that characterizes the variation in the distribution.

Other questions in which spread is the issue are

- What is the variability in timber production among national forests?

- What is the variation among the states in food stamp participation rates?
- What is the spread in asset value among failed savings and loan institutions?

In each of these examples, we are addressing the generic question, How much spread (or variation in the response variable) is there among the cases? (See table 1.3.)

Even when spread is not the center of attention, it is an important concept in data analysis and should be reported when a set of data are described. Whenever evaluators give information about the central tendency of a distribution, they should also describe the spread.

Measures of the Spread of a Distribution

There is a variety of statistics for gauging the spread of a distribution. Some measures should be used only with interval-ratio measurement while others are appropriate for nominal or ordinal data. Table 3.1 summarizes the characteristics of four particular measures.

Table 3.1: Measures of Spread

Measurement level	Use of measure			
	Index of dispersion	Range	Interquartile range	Standard deviation
Nominal	Yes	No	No	No
Ordinal	Sometimes	Sometimes	Yes	No
Interval-ratio	No	Yes	Yes	Yes

The index of dispersion is a measure of spread for nominal or ordinal variables. With such variables, each case falls into one of a number of categories. The index shows the extent to which cases are

bunched up in one or a few categories rather than being well spread out among the available categories.

The calculation of the index is based upon the concept of unique pairs of cases. Suppose, for example, we want to know the spread for gender, a nominal variable. Assume a batch of 8 cases, 3 females and 5 males. Each of the 3 females could be paired with each of the 5 males to yield 15 unique pairs (3×5).

The index is a ratio in which the numerator is the number of unique pairs (15 in the example) that can be created given the observed number of cases ($n = 8$ in the example). The denominator of the ratio is the maximum number of unique pairs of cases that can be created with n cases. The maximum occurs when the cases are evenly divided among the available categories.

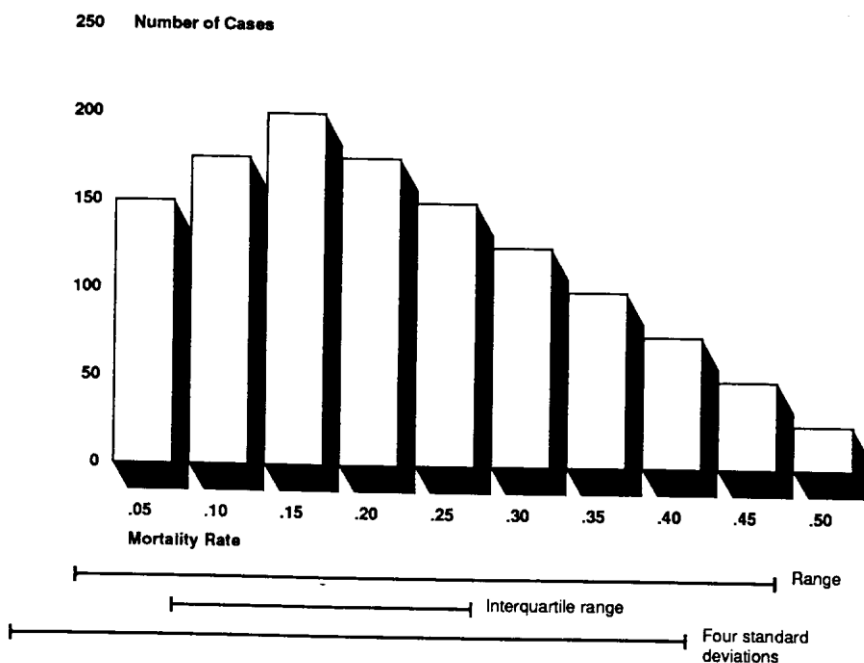
The maximum number of unique pairs (for $n = 8$) would occur if the batch included 4 females and 4 males (the 8 cases evenly divided among the two categories). Under this condition, 16 unique pairs (4×4) could be formed. The index of dispersion for the example would thus be $15/16 = .94$. Although this example illustrates the concept of the index, the calculation of the index becomes more tedious as the number of cases and the number of categories increase. Loether and McTavish (1988) give a computational formula and a computer program for the index of dispersion.

As the cases become more spread out among the available categories, the index of dispersion increases in value. The index of dispersion can be as large as 1, when the categories have equal numbers of cases, and as small as 0, when all cases are in one category.

The range is a commonly used measure of spread when a variable is measured at least at the ordinal level. The range is the difference between the largest and smallest observations in the distribution. Because the range is based solely on the extreme values, it is a crude measure that is very sensitive to sample size and to outliers. The effect of an outlier is shown by the two distributions we considered in chapter 2: (1) 1,4,4,5,7,7,8,8,9, and (2) 1,4,4,5,7,7,8,8,542. The range for the first distribution is 8, and for the second it is 541. The huge difference is attributable to the presence of an outlier in the second distribution.

A range of 0 means there is no variation in the cases, but unlike the index of dispersion, the range has no upper limit. The range is not used with nominal variables because the measure makes sense only when cases are ordered. To illustrate the measure, the distribution of hospital mortality rates, is reproduced in figure 3.2. Inspection of the data showed that the minimum rate was .025 and the maximum was .475, so the range is .45.

Figure 3.2: Spread of a Distribution



Another measure of spread, the interquartile range, is the difference between the two points in a distribution that bracket the middle 50 percent of the cases. These two points are called the 1st and 3rd quartiles and, in effect, cut the upper and lower 25 percent of the cases from the range. The more closely the cases are bunched together, the smaller will be the value of the interquartile range. Like the range, the interquartile range requires at least an ordinal level of measurement, but by discounting

extreme cases, it is not subject to criticism for being inappropriately sensitive to outliers. In the hospital mortality example, the 1st quartile is .075 and the 3rd quartile is .275 so the interquartile range is the difference, .2.

A fourth measure of spread, one often used with interval-ratio data, is the standard deviation. It is the square root of the average of the squares of the deviations of each case from the mean. As with the preceding measures, the standard deviation is 0 when there is no variation among the cases. It has no upper limit, however. For the distribution of hospital mortality rate, the standard deviation is .12 but note, from figure 3.2, that the distribution is somewhat asymmetric, so this measure of spread is apt to be misleading. The four-standard-deviation band shown in figure 3.2 is .48 units wide and centered on the sample mean of .19.¹

One way of interpreting or explaining the spread of a distribution (for ordinal or higher variables) is to look at the proportion of cases “covered” by a measure of dispersion. To do this, we think of a spread measure as a band having a lower value and an upper value and then imagine that band superimposed on the distribution of cases. A certain proportion of the cases have observations larger than the lower value of the band and less than the upper value; those cases are thus covered by the spread measure. For the range, the lower value is the smallest observation among all cases and the upper value is the largest observation (see figure 3.2, based upon 1,225 cases). Then 100 percent of the cases are covered by the range.

¹Expressing the spread as a band of four standard deviations is a common but not unique practice. Any multiple of standard deviations would be acceptable but two, four, and six are commonplace.

Chapter 3

Determining the Spread of a Distribution

Likewise, we know that when the interquartile range is used, 50 percent of the cases are always covered. The situation with the standard deviation is more complex but ultimately, in terms of inferential statistics, more useful.

When we use the standard deviation as the measure of spread, we can define the width of the band in an infinite number of ways but only two or three are commonly used. One possibility is to define the lower value of the band as the mean minus one standard deviation and the upper value as the mean plus one standard deviation. In other words, this band is two standard deviations wide (and centered on the mean). We could then simply count the cases in the batch that are covered by the band. However, it is important to realize that the number of cases can vary from study to study. For example, 53 percent of the cases might be covered in one study, to pick an arbitrary figure, and 66 percent in another. Just how many depends upon the shape of the distribution. So, unlike the situation with the range or the interquartile range, the measure by itself does not imply that a specified proportion of cases will be covered by a band that is two standard deviations wide. Thus if we know only the width of the band, we may have difficulty interpreting the meaning of the measure. Other bands could be defined as four standard deviations wide or any other multiple of the basic measure, a standard deviation.²

We can obtain some idea of the effect of distribution shape on the interpretation of the standard deviation

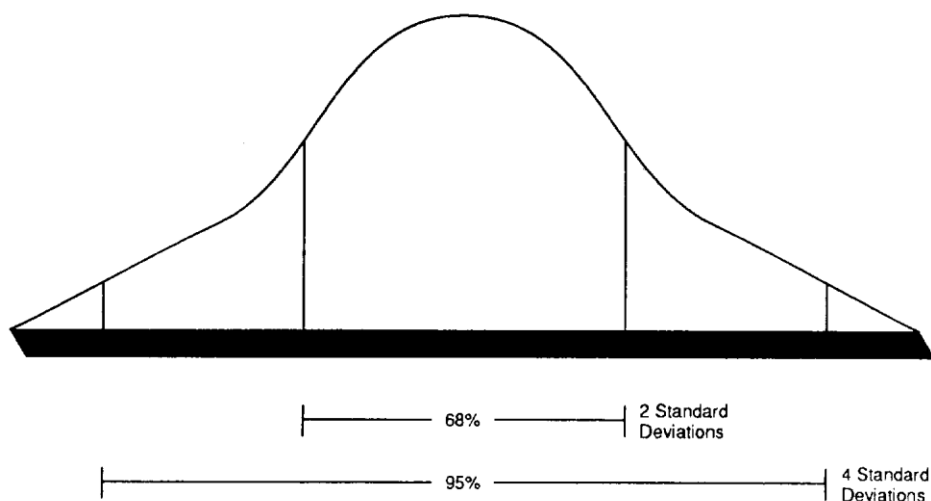
²The term “standard deviation” is sometimes misunderstood to be implying some substantive meaning to the amount of variation—that the variation is a large amount or a small amount. The measure by itself does not convey such information, and after we have computed a standard deviation, we still have to decide, on the basis of nonstatistical information, whether the variation is “large” or not.

by considering three situations. We may believe that the distribution is (1) close to a theoretical curve called the normal distribution (the familiar bell-shaped curve, (2) has a single mode and is approximately symmetric (but not necessarily normal in shape), or (3) of unknown or “irregular” shape.³ For this example, we define the band to be four standard deviations wide (that is, two standard deviations on either side of the mean).

When the distribution of a batch is close to a normal distribution, statistical theory permits us to say that approximately 95 percent of the cases will be covered by the four-standard-deviation band. (See figure 3.3.) However, if we know only that the distribution is unimodal and symmetric, theory lets us say that, at minimum, 89 percent of the cases will be covered. If the distribution is multimodal or asymmetric or if we simply do not know its shape, we can make a weaker statement that applies to any distribution: that, at minimum, 75 percent of the cases will be covered by the four-standard-deviation band.

³The name for the set of theoretical distributions called “normal” is unfortunate in that it seems to imply that distributions that have this form are “to be expected.” While many real-world distributions are indeed close to a normal (or Gaussian) distribution in shape, many others are not.

Figure 3.3: Spread in a Normal Distribution



From this example, it should be evident that care must be taken when using the standard deviation to describe the spread of a batch of cases. The common interpretation that a four-standard-deviation band covers about 95 percent of the cases is true only if the distribution is approximately normal.

One GAO example of describing the spread of a distribution comes from a report on Bureau of the Census methods for estimating the value of noncash benefits to poor families (U.S. General Accounting Office, 1987b). Variation in the amount of noncash benefits was described in terms of both the range and the standard deviation. In a study of homeless children and youths (U.S. General Accounting Office,

1989a), GAO evaluators asked shelter providers, advocates for the homeless, and government officials to estimate the proportion of the homeless persons, in a county, that seek shelter in a variety of settings (for example, churches, formal shelters, and public places). The responses were summarized by reporting medians and by the first and third quartiles (from which the interquartile range can be computed).

Analyzing and Reporting Spread

To analyze the spread of a nominal variable, it is probably best just to develop a table or a histogram that shows the frequency of cases for each category of the variable. The calculation of a single measure, such as the index of dispersion, is not common but can be done.

For describing the spread of an ordinal variable, tables or histograms are useful, but the choice of a single measure is problematic. The index of dispersion is a possibility, but it does not take advantage of the known information about the order of the categories. Range, interquartile range, and standard deviation are all based on interval or ratio measurement. When a single measure is used, the best choice is often the interquartile range.

With an interval-ratio variable, graphic analysis of the spread is always advisable even if only a single measure is ultimately reported. The standard deviation is a commonly used measure but, as noted above, may be difficult to interpret if it cannot be shown that the cases have approximately a normal distribution.⁴ Consequently, the interquartile range

⁴A possible approach with a variable that does not have a normal distribution is to change the scale of the variable so that the shape does approximate the normal. See Velleman and Hoaglin (1981) for some examples; they refer to the process of changing the scale as “re-expression,” but “transformation” of the variables is a more common term.

may be a good alternative to the standard deviation when the distribution is questionable.

With respect to reporting data, a general principle applies: whenever central tendency is reported, spread should be reported too. There are two main reasons for this. The first is that a key study question may ask about the variability among cases. In such instances, the mean should be reported but the real issue pertains to the spread.

The second reason for describing the spread of a distribution, which applies even when the study question focuses on central tendency, is that knowledge of variation among individual cases tells us the extent to which an action based on the central tendency is likely to be on the mark. The point is that government action based upon the central tendency may be appropriate if the spread of cases is small, but if the spread is large, several different actions may be warranted to take account of the great variety among the cases. For example, policymakers might conclude that the mean mortality rate among hospitals is satisfactory and, given central tendency alone, might decide that no action is needed. If there is little spread among hospitals with respect to mortality rates, then taking no action may be appropriate. But if the spread is wide, then maybe hospitals with low rates should be studied to see what lessons can be learned from them and perhaps hospitals with extremely high rates should be looked at closely to see if improvements can be made.

Determining Association Among Variables

Many questions GAO addresses deal with associations among variables:

- Do 12th grade students in high-spending school districts learn more than students in low-spending districts?
- Are different procedures for monitoring thrift institutions associated with different rates for correctly predicting institution failure?
- Is there a relationship between geographical area and whether farm crop prices are affected by price supports?
- Are homeowners' attitudes about energy conservation related to their income level?
- Are homeowners' appliance-purchasing decisions associated with government information campaigns aimed at reducing energy consumption?

Recalling table 1.3, these examples illustrate the third generic question, To what extent are two or more variables associated? An answer to the first question would reveal, for example, whether high achievement levels tend to be found in higher-spending districts and low achievement levels in lower-spending districts (a positive association), or vice versa (a negative association).

What Is an Association Among Variables?

Just what do we mean by an association among variables?¹ The simplest case is that involving two variables, say homeowners' attitudes about energy conservation and income level. Imagine a data sheet as in table 4.1 representing the results of interviews with 341 homeowners. For these hypothetical data, we have adopted the following coding scheme: attitude toward energy conservation (indifferent = 1, somewhat positive = 2, positive = 3); family income level (low = 1, medium = 2, high = 3).

¹The term "relationship" is equivalent to "association."

**Table 4.1: Data Sheet
With Two Variables**

Case	Attitude toward energy conservation	Family income level
1	3	2
2	3	1
3	1	3
341	2	2

To say that there is an association between the variables is to say that there is a particular pattern in the observations. Perhaps homeowners who respond that their attitude toward energy conservation is positive tend to report that they have low income and homeowners who respond that they are indifferent toward conservation tend to have high income. The pattern is that the cases vary together on the two variables of interest. Usually the relationship does not hold for every case but there is a tendency for it to occur.

The trouble with a data sheet like this is that it is usually not easy to perceive an association between the two variables. Evaluators need a way to summarize the data. One common way, with nominal or ordinal data, is to use a cross-tabular display as in table 4.2. The numbers in the cells of the table indicate the number of homeowners who responded to each possible combination of attitude and income level.

Table 4.2:
Cross-Tabulation of Two
Ordinal Variables

Attitude toward energy conservation	Family income level			Total
	Low	Medium	High	
Indifferent	27	37	56	120
Somewhat positive	35	39	41	115
Positive	43	33	30	106
Total	105	109	127	341

Notice that the information in table 4.2 is an elaboration of the distribution of 341 homeowners shown in table 1.2. Reading down the total column in table 4.2 gives the distribution of the homeowners with respect to the attitude variable (the same as in table 1.2). In a two-variable table, this distribution is called a marginal distribution; it presents information on only one variable. The last row in table 4.2 (not including the grand total, 341) also gives a marginal distribution—that for the income variable.

There is much more information in table 4.2. If we look down the numbers in the low-income column only, we are looking at the distribution of attitude toward energy conservation for only low-income households. Or, if we look across the indifferent row, we are looking at the distribution of income levels for indifferent households. The distribution of one variable for a given value of the other variable is called a conditional distribution. Four other conditional distributions (for households with medium income, high income, somewhat positive attitudes, and positive attitudes) are displayed in table 4.2, which in its entirety portrays a bivariate distribution.

The new table compresses the data, from 682 cells in the data sheet of table 4.1 to 16, and again we can look for patterns in the data. In effect, we are trying to compare distributions (for example, across

low-income, medium-income, and high-income households) and if we find that the distributions are different (across income levels, for example) we will conclude that attitude toward energy conservation is associated with income level. Specifically, households with high income tend to be less positive than low-income households. But the comparisons are difficult because the number of households in the categories (for example, low-income and medium-income) are not equal, as we can observe from the row and column totals.

The next step in trying to understand the data is to convert the numbers in table 4.2 to percentages. That will eliminate the effects of different numbers of households in different categories. There are three ways to make the conversion: (1) make each number in a row a percentage of the row total, (2) make each number in a column a percentage of a column total, or (3) make each number in the table a percentage of the batch total, 341. (Computer programs may readily compute all three variations.) In table 4.3, we have chosen the second way. Now we can see much more clearly how the distributions compare for different income levels.

Table 4.3: Percentaged Cross-Tabulation of Two Ordinal Variables

Attitude toward energy conservation	Family income level			Total
	Low	Medium	High	
Indifferent	26	34	44	35
Somewhat positive	33	36	32	34
Positive	41	30	24	31
Total	100	100	100	100

And we could go on and look at the other ways of computing percentages. But even with all three displays, it still may not be easy to grasp the extent of an association, much less readily communicate its

extent to another person. Therefore, we often want to go beyond tabular displays and seek a number, a measure of association, to summarize the association. Such a measure can be used to characterize the extent of the relationship and, often, the direction of the association, except for nominal variables. We may sometimes use more than one measure to observe different facets of an association. Although this example involves two ordinal variables, the notion of an association is similar for other combinations of measurement levels.

Measures of Association Between Two Variables

A measure of association between variables is calculated from a batch of observations, so it is another descriptive statistic. Several measures of association are available to choose from, depending upon the measurement level of the variables and exactly how association is defined. For illustrative purposes, we mention four from the whole class of statistics sometimes used for indicating association: gamma, lambda, the Pearson product-moment correlation coefficient, and the regression coefficient.

Ordinal Variables: Gamma

When we have two ordinal variables, as in the energy conservation example, gamma is a common statistic used to characterize an association. This indicator can range in value from -1 to $+1$, indicating perfect negative association and perfect positive association, respectively. When the value of gamma is near zero, there is little or no evident association between the two variables. Gamma is readily produced by available statistical programs, and it can be computed by hand from a table like table 4.2, but the calculation, sketched out below, is rather laborious. For our hypothetical data set, gamma is found to be $-.24$.

The computed value of gamma indicates that the association between family income level and attitude toward energy conservation is negative but that the extent of the association is modest. One way to interpret this result is that if we are trying to predict a family's attitude toward energy conservation, we will be more accurate (but not much more) if we use knowledge of its income level in making the prediction. The gamma statistic is based upon a comparison of the errors in predicting the value of one variable (for example, family's attitude toward conservation) with and without knowing the value of another variable (family income). This idea is expressed in the following formula: $\text{gamma} = (\text{prediction errors not knowing income} - \text{prediction errors knowing income}) / \text{prediction errors not knowing income}$.

The calculation of gamma involves using the information in table 4.2 to determine the number of prediction errors for each of two situations, with and without knowing income. The formula above is actually quite general and applies to a number of measures of association, referred to as PRE (proportionate reduction in error) measures. The more general formulation (Loether and McTavish, 1988) is $\text{PRE measure} = \text{reduction in errors with more information} / \text{original amount of error}$. PRE measures vary, depending upon the definition of prediction error.

Nominal Variables:
Lambda

With two nominal variables, the idea of an association is similar to that between ordinal variables but the approach to determining the extent of the association is a little different. This is so because, according to definition, the attributes of a nominal variable are not ordered. The consequences can be seen by looking at another cross-tabulation.

Suppose we have data with which to answer the question about the association between whether the prices farmers receive are affected by government crop supports and the region of the country in which they live. Then the variables and attributes might be as follows: crop supports (yes, no); region of the country (Northeast, Southeast, Midwest, Southwest, Northwest). Some hypothetical data for these variables are displayed in table 4.4.

Table 4.4:
Cross-Tabulation of Two
Nominal Variables

Region	Region Prices affected by crop		Total
	Yes	No	
Northeast	322	672	994
Southeast	473	287	760
Midwest	366	382	748
Southwest	306	297	748
Northwest	342	312	654
Total	1,809	1,950	3,759

If we start to look for a pattern in this cross-tabulation, we have to be careful because the order in which the regions are listed is arbitrary. We could just as well have listed them as Southwest, Northeast, Northwest, Midwest, and Southeast or in any other sequence. Therefore, the pattern we are looking for cannot depend upon the sequence as it does with ordinal variables.

Lambda is a measure of association between two nominal variables. It varies from 0, indicating no association, to 1, indicating perfect association.² The calculation of lambda, which is another PRE measure

²A definition of perfect association is beyond the scope of this paper. Different measures of association sometimes imply different notions of perfect association.

like gamma, involves the use of the mode as a basis for computing prediction errors. For the crop support example, the computed value of lambda is .08.³ This small value indicates that there is not a very large association between crop-support effects and region of the country.

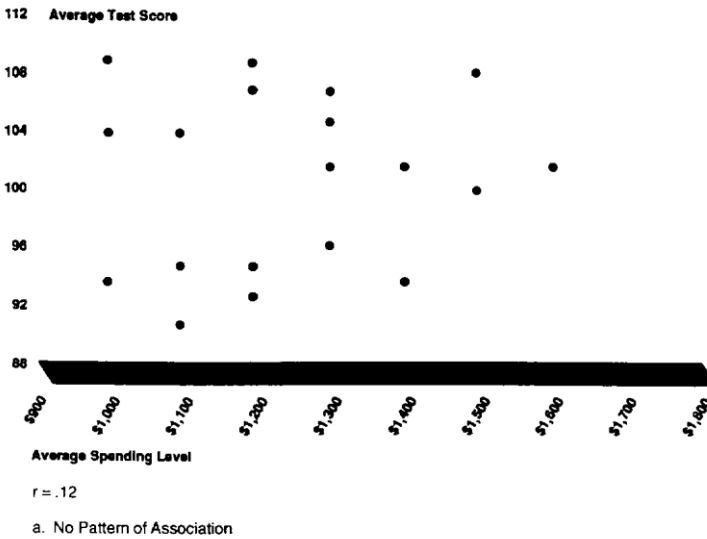
**Interval-Ratio
Variables:
Correlation and
Regression
Coefficients**

A Pearson product-moment correlation coefficient is a measure of linear association between two interval-ratio variables.⁴ The measure, usually symbolized by the letter r , varies from -1 to $+1$, with 0 indicating no linear association. The square of the correlation coefficient is another PRE measure of association.

³There are actually three ways to compute lambda. The numerical value here is the symmetric lambda. There is some discussion of symmetric and asymmetric measures of association later in this paper.

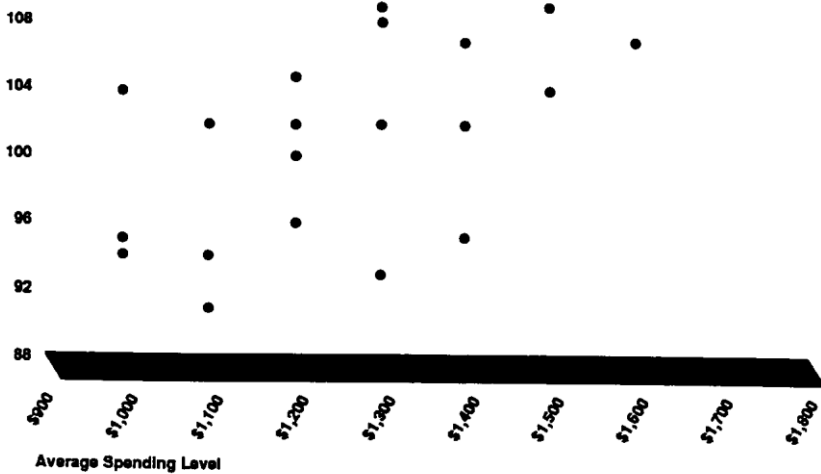
⁴The word "correlation" is sometimes used in a nonspecific way as a synonym for "association." Here, however, the Pearson product-moment correlation coefficient is a measure of linear association produced by a specific set of calculations on a batch of data. It is necessary to specify linear because if the association is nonlinear, the two variables might have a strong association but the correlation coefficient could be small or even zero. This potential problem is another good reason for displaying the data graphically, which can then be inspected for nonlinearity. For a relationship that is not linear, another measure of association, called "eta," can be used instead of the Pearson coefficient (Loether and McTavish, 1988).

Scatter Plots for Spending Level and Test Scores



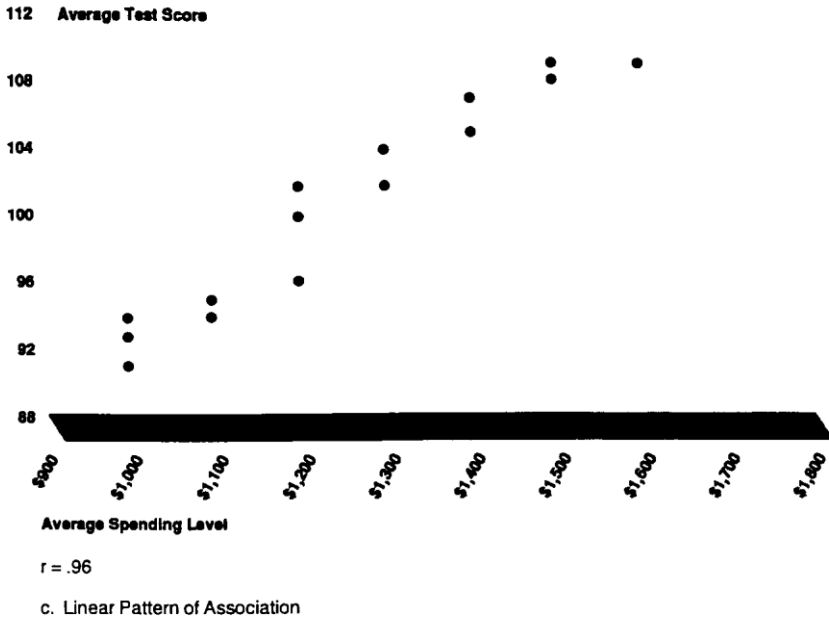
The Pearson product-moment correlation coefficient can be illustrated by considering the question about the association between students' achievement level in the 12th grade and school district spending level, regarding both variables as measured at the interval-ratio level. Such data are often displayed in a scatter plot, an especially revealing way to look at the association between two variables measured at the interval-ratio level. Figure 4.1 shows three scatter plots for three sets of hypothetical data on two variables: average test scores for 12th graders in a school district and the per capita spending level in the district. Each data point represents two numbers: a districtwide test score and a spending level.

112 Average Test Score



$r = .53$

b. Semi-Patterned Association



In figure 4.1a, which shows essentially no pattern in the scatter of points, the correlation coefficient is .12. In figure 4.1b, the points are still widely scattered but the pattern is clear—a tendency for high test scores to be associated with high spending levels and vice versa; the correlation coefficient is .53. And finally, in figure 4.1c the linear pattern is quite pronounced and the correlation coefficient is .96.

The regression coefficient is another widely used measure of association between two interval-ratio variables and it can be used to introduce the idea of

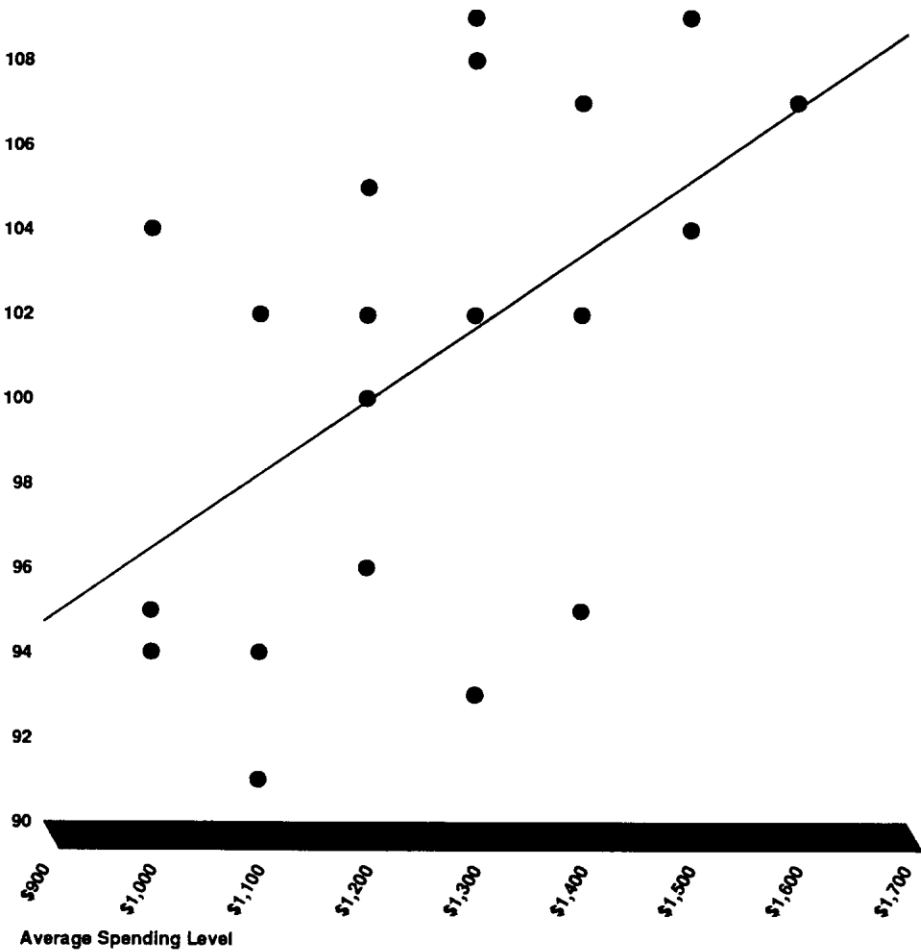
an asymmetric measure. First, we use the scatter plot data from figure 4.1b and replot them in figure 4.2. Using the set of data represented by the scatter plot, we can use the method of regression analysis to “regress Y on X,” which tells us where to position a line through the scatter plot.⁵ How the analysis works is not important here, but the interpretation of the line as a measure of association is. The slope of the line is numerically equal to the amount of change in the Y variable per 1 unit change in the X variable. The slope is the regression coefficient, an asymmetric measure of association between the two variables. Unlike many other commonly used measures, the regression coefficient is not limited to the interval from -1 and $+1$.⁶ The regression coefficient for the data displayed in figure 4.2 is 1.76, indicating that a \$100 change in spending level is associated with a 1.76 change in test scores.

⁵Regression analysis is not covered in this paper. For extensive treatments, see Draper and Smith, 1981, and Pedhazur, 1982.

⁶The regression coefficient is closely related to the Pearson product moment correlation. In fact, when the observed variables are transformed to so-called z-scores, by subtracting the mean from each observed value of a variable and dividing the difference by the standard deviation of the variable, the regression coefficient of the transformed variables is equal to the correlation coefficient.

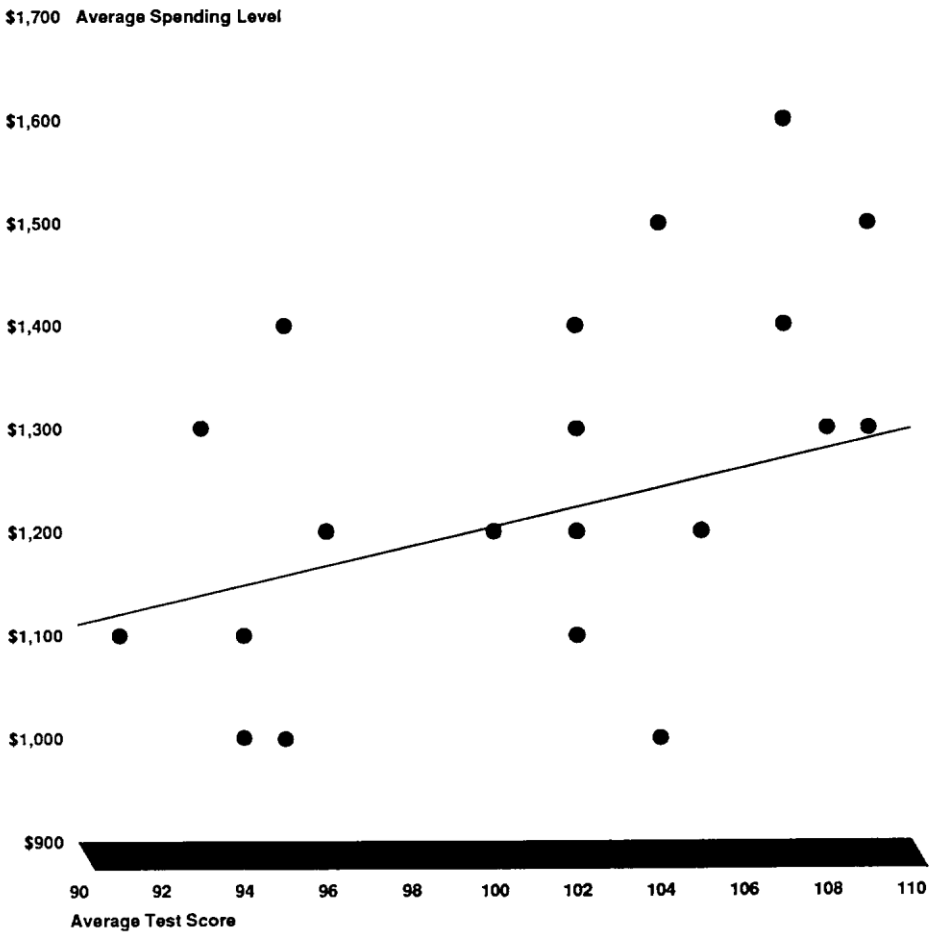
Regression of Test Scores on Spending Level

110 Average Test Score



Why the regression coefficient is asymmetric can be understood if we turn the scatter plot around, as in figure 4.3, so that X is on the vertical axis and Y is on the horizontal. The pattern of points is a little different now, and if we reverse the roles of the X and Y variables in the regression procedure (that is, “regress X on Y”), the resulting line will have a different slope. Consequently the Y-on-X regression coefficient is different from the X-on-Y coefficient and that is why the measure is said to be asymmetric. Measures of association in which the roles of the X and Y variables can be interchanged in the calculation procedures without affecting the measure are said to be symmetric and measures in which the interchange produces different results are asymmetric.

Figure 4.3: Regression of Spending Level on Test Scores



When we use asymmetric measures of association, we also use special language to characterize the roles of the variables. Or, put in a more direct way, if we take the view that the variables are playing different roles, we give them names indicative of the roles. One is called the dependent variable and the other is the independent variable. The language is applied to two kinds of application: (1) when we are trying to establish that the independent variable causes changes in the dependent one and (2) when we are trying to use the independent variable to predict the dependent one, without necessarily supposing the association is causal. In either case, the dependent variable in some sense depends upon the independent one. Graphically, the convention is to plot the dependent variable along the vertical axis and the independent variable along the horizontal axis.

Whether evaluators should use an asymmetric measure of association or a symmetric one depends upon the application. If there is no reason to label variables as dependent and independent, then they should use a symmetric measure. But when they are predicting one variable from another or believe that one has a causal effect on the other, an asymmetric measure is preferred.

In each of the foregoing examples, both variables were measured at the same level. That will not always be the case. One common circumstance in which the variables are at different levels is discussed in a section below, entitled “The Comparison of Groups.”

Examples

An example of a measure of association between ordinal variables comes from a GAO report on the use of medical devices in hospitals (U.S. General Accounting Office, 1986). In reporting on the association between the seriousness of a device

problem and hospitals' actions to contact the manufacturer or some other party outside the hospital, the evaluators displayed the results in cross-tabular array and summarized them using a symmetric measure.

In a study of election procedures (U.S. General Accounting Office, 1990d), some of the major findings were reported as a series of correlation coefficients that showed the association between voter turnout and numerous factors characterizing absentee ballot rules and voter information activities. The same study used a regression coefficient to show the association between voter turnout and the registration deadline, expressed as number of days before the election.

The Comparison of Groups

A situation of special interest arises when evaluators want to compare two groups on some variable to see if they are different. For example, suppose the evaluators want to compare government benefits received by farmers who live east of the Mississippi to those who live west of the Mississippi. Questions about the difference between two groups are very common. In this instance, it would probably be best to answer the question by computing the mean benefits for each group and looking at the difference.

Equivalently, the comparison between these two groups of cases can be seen as a measure of association. With government benefits measured at the interval-ratio level (in dollars) and region of the country measured at the nominal level (for example, 0 for East and 1 for West), we can compute a measure of association called the point biserial correlation between benefits and region.⁷ If we then multiply this correlation by the standard deviation of benefits and

⁷The point biserial correlation is analogous to the Pearson product-moment correlation that applies when both variables are measured at the interval-ratio level.

divide by the standard deviation of region, we will get the difference between the means of the two groups. The same result would be obtained if we regressed benefits on region; the regression coefficient is equal to the difference between the means of the two groups. We thus have three different, but statistically equivalent, methods of comparing the two groups: (1) computing the difference between means of the groups, (2) computing the point biserial correlation (and then adjusting it), and (3) computing the regression coefficient.

The point is that when evaluators compare two groups, they are examining the extent of association between two variables: one is group membership and the other is the response variable, the characteristic on which the groups might differ. Such comparisons are the main method for evaluating the effect of a program. For example, a question might be: What is the effect of the Special Supplemental Food Program for Women, Infants, and Children (WIC) on birthweight? The answer is partly to be found in the association, if any, between group membership (program participation or not) and birthweight.

Knowing the association is only part of the answer, however, because the question about effect is about the causal association between program participation and birthweight. As we show in chapter 6, the existence of an association is one of three conditions necessary to establish causality.

A comparison of means is but one among many ways in which it might be appropriate to compare two groups. Other possibilities include the comparison of (1) medians, (2) proportions, and (3) distributions. For example, if two groups are being compared on an ordinal variable and the distribution is highly asymmetric, then an analysis of the medians may be

preferable to an analysis of the means. Or, as noted in chapter 3, sometimes the question the evaluators are attempting to answer is focused on the spread of a distribution and so they might be interested in comparing a measure of spread in two groups. For the hospital mortality study, we could compare the spread of mortality rates between two categories of hospitals, say teaching and nonteaching ones.

Statistical methods for comparing groups are important to GAO in at least three situations: (1) comparison of the characteristics of populations (for example, farmers in the eastern part of the country with those in the western), (2) determination of program effects (for example, the WIC program), and (3) the comparison of processes (for example, different ways to monitor thrift institutions). The questions that arise from these situations lead to a variety of data analysis methods. Factors that determine an appropriate data analysis methodology include (1) the number of groups to be compared, (2) how cases for the groups were selected, (3) the measurement level of the variables, (4) the shape of the distributions, and (5) the type of comparison (measure of central tendency, measure of spread, and soon). A further complexity is that, when sampling, evaluators need to know if the observed difference between groups is real or most likely stems from sampling fluctuation. For making that determination, the methods of statistical inference are required.

A study of changes to the program called Aid to Families with Dependent Children illustrates the use of group comparisons on factors such as employment status to draw conclusions about effects of the changes (U.S. General Accounting Office, 1985). In another example, two groups of farmers, ones who specialized in a few crops and ones who diversified,

were compared on agricultural practices (U.S. General Accounting Office, 1990a).

Analyzing and
Reporting the
Association
Between
Variables

Answering a question about the association between two variables really involves four subquestions: Does an association exist? What is the extent of the association? What is the direction of the association? What is the nature of the association? Analysis of a batch of data to answer these questions usually involves the production of tabular or graphic displays as well as the calculation of measures of association.

With nominal or ordinal data presented in tabular form, evaluators can check for the existence of an association by inspection of the tables. If the conditional distributions are identical or nearly so, they can conclude that there is no association. Table 4.2 illustrates a data set for which an association exists between income level and attitude toward energy conservation. Table 4.5 shows another set of 341 cases—one in which there is virtually no association. The marginal distributions are the same for tables 4.2 and 4.5, so the pattern of observations can change only in the nine interior cells.

Table 4.5: Two Ordinal
Variables Showing No
Association

Attitude toward energy conservation	Family income level			Total
	Low	Medium	High	
Indifferent	37	38	45	120
Somewhat positive	35	37	43	115
Positive	33	34	39	106
Total	105	109	127	341

Most bivariate data show the existence of association. The question is really whether the association is large

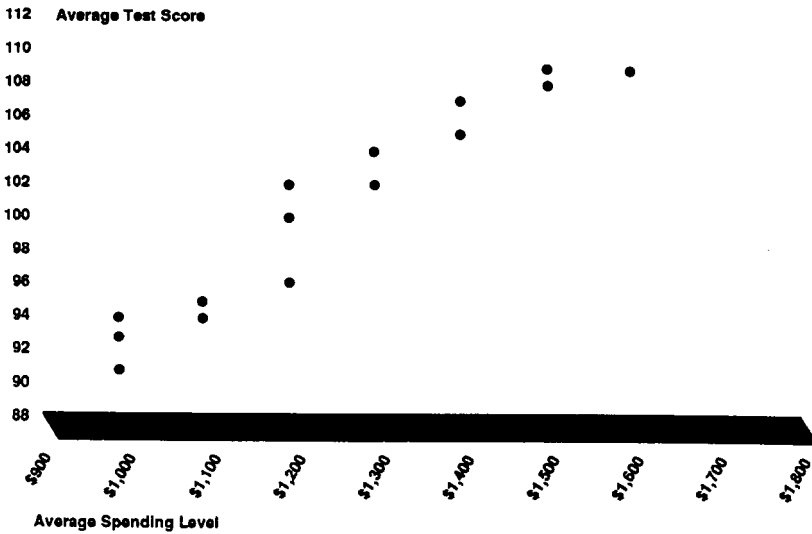
enough to be important.⁸ A measure of association is calculated to help answer this question, and evaluators must make a judgment about importance, using the context of the question as a guide.

The direction of an association is also given by a measure of association unless the variables are nominal, in which case the direction is not meaningful. Most measures are defined so that a negative value indicates that as one variable increases the other decreases and that a positive value indicates that the variables increase or decrease together.

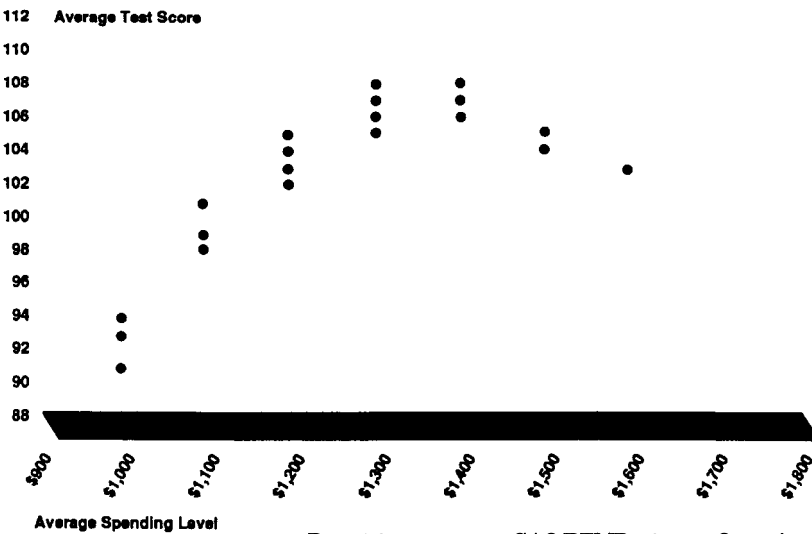
While the existence, extent, and direction of an association can be revealed by a measure of association, determining the nature of the association requires other methods. Usually it is done by inspecting the tabular or graphic display of a bivariate distribution. For example, a scatter plot will show if the association is approximately linear, a constant amount of change in one variable being associated with a constant amount in the other variable, as in figure 4.4a. However, the scatter plot may show that the association is nonlinear, as in figure 4.4b. Interpretations of the data are usually easier if the data are linear and, of course, interpolations and extrapolations are more straightforward.

⁸If we are trying to draw conclusions about a population from a probability sample, then we must additionally be concerned about whether what seems to be an association really stems from sampling fluctuation. The data analysis then involves inferential statistics.

Figure 4.4: Linear and Nonlinear Associations



a. Linear Association



b. Nonlinear Association

In comparisons between groups of cases, regression analysis is an important tool when the dependent variable is interval-ratio. When the assumptions necessary for regression are not satisfied, other techniques are necessary.⁹

Overall, there are many analysis choices. Evaluators can always find the extent of the association, if any, between the variables and, unless one or both the variables are measured at the nominal level, they can also determine the direction of the association. The appropriateness of a given procedure depends upon the measurement level of the variables and the definition of association believed best for the circumstances. It is also wise to display the data in a table or a graph as a way to understand the form of the association.

How much information from the analysis should be included in a report? The answer depends on how strongly the conclusions are based upon the association that has been determined. If the relationship between the two variables is crucial, then probably both measures of association and tabular or graphic displays should be presented. Otherwise, reporting only the measures will probably suffice. In either case, evaluators should be clear about the level of measurement assumed and analysis methods used.

⁹The assumptions are not very stringent for descriptive statistics but may be problematic for inferential statistics.

Estimating Population Parameters

Many questions that GAO seeks to answer are about relatively large populations of persons, things, or events. Examples are

- What is the average student loan balance owed by college students?
- Among households eligible for food stamps, what proportion receive them?
- How much hazardous waste is produced in the nation annually and how much variation is there among individual generators?
- What is the relationship between the receipt of Medicaid benefits and size of household?

In chapters 2, 3, and 4, we focused on descriptive statistics—ways to answer questions about just those cases for which we had data. We now consider inferential statistics—methods for answering questions about cases for which we do not have observations. The procedures involve using data from a sample of cases to infer conclusions about the population of which the sample is a part.

The shift to inferential statistics is necessary when evaluators want to know about large populations but, for practical reasons, do not try to get information from every member of such populations. The most obvious obstacle to collecting data on many cases is cost, but other factors such as deadlines for producing results may play a role.

To generalize findings from a sample of cases to the larger population, not just any sample of cases will do—a probability sample is required. Random processes for drawing probability samples are detailed in the transfer paper in this series entitled Using Statistical Sampling.¹ Under such methods,

¹Probability sampling is sometimes called statistical sampling or scientific sampling.

each member of a population has a known, nonzero probability of being drawn.

The methods collectively called inferential statistics are based upon the laws of probability and require samples drawn by a random process. Attempts to draw conclusions about populations based upon nonprobability samples are usually not very persuasive, so we do not consider them here.

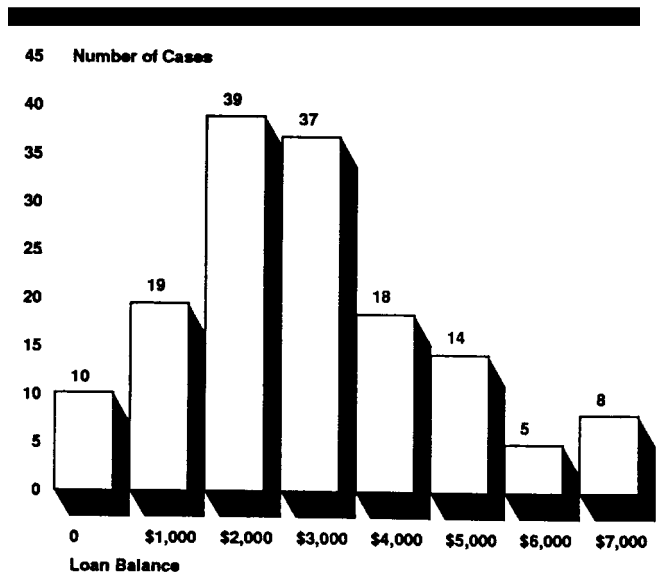
From the perspective of inferential statistics, the illustrative questions above need two-part answers: a point estimate of a parameter that describes the population and an interval estimate of the parameter. (Other forms of statistical inference, such as hypothesis testing, are appropriate to other kinds of questions. They are not covered in this paper.) Full understanding of inferential statistical statements requires a thorough knowledge of probability, the development of which is beyond the scope of this introductory paper. For our brief treatment here, we use the concept of the histogram and illustrate how probability comes into play through sampling distributions.

Some notions discussed in earlier chapters, involving data on all cases in a batch, are extended in this chapter to show how statistics computed from a probability sample of cases are used to estimate parameters such as the central tendency of a population (see chapter 2). The notable difference between describing a batch, using statistics from all cases in the batch, and describing a population, using statistics from a probability sample of the population, is that we will necessarily be somewhat uncertain in describing a population. However, the data analysis methods for inferential statistics allow us to be precise about the degree of uncertainty.

Histograms and Probability Distributions

A key concept in statistical inference is the sampling distribution. The histogram, which was introduced in chapter 1, is a way of displaying a distribution, so we begin there. Expanding on the first example from chapter 1, suppose that instead of information on a batch of 15 college students, we have collected information on loan balances from 150 students. If we round numbers to the nearest \$1,000 for ease of computation and display, our observations produce the distribution of loan balances shown in figure 5.1. For example, the height of the third bar corresponds to the number of students who reported loan balances between \$1,500 and \$2,499. The distribution is somewhat asymmetrical and has a mean of \$2,907.

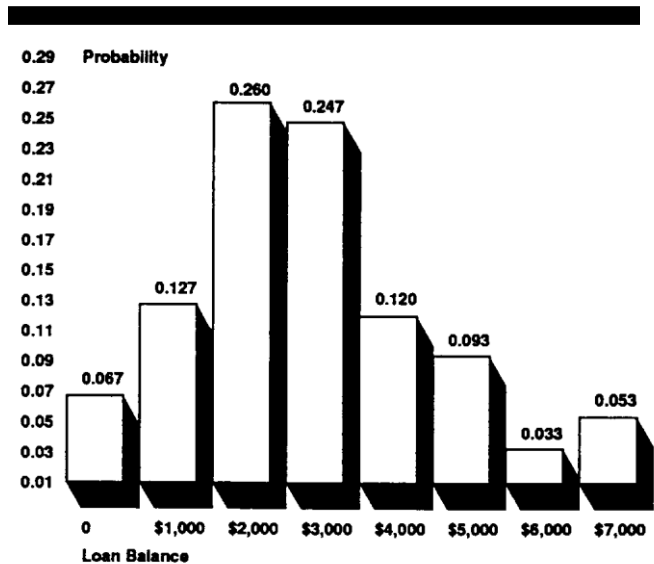
Figure 5.1: Frequency Distribution of Loan Balances



Probability is a numerical way of expressing the likelihood that a particular outcome, among a set of possibilities, will occur. Suppose that we do not have access to the responses from individual students in the survey but that we want to use the distribution in figure 5.1 to make a wager on whether a student to be selected at random from this sample of 150 will have a loan balance between \$1,500 and \$2,500. To make a reasonable bet, we need to know the probability that a particular outcome—a loan value between \$1,500 and \$2,500—will be reported when we make a phone call to the student. The information we need is in the figure but the answer will be more evident if we make a slight change in the display.

We can describe the students' use of loans in probability terms if we convert the frequency distribution to a probability distribution. The frequency distribution shows the number of students who reported each possible outcome (that is, loan balances between \$1,500 and \$2,500 and so on). We can present the same information in terms of percentages by dividing the number of students reporting each outcome (the height of a bar) by the total number in the sample (150). The percentages, expressed in decimal form, can be interpreted as probabilities and are displayed in figure 5.2.

**Figure 5.2: Probability
Distribution of Loan
Balances**



We now have a probability distribution for the loan balance variable for the sample of 150 students. Picking the outcome we want to make a wager on, we can say that the probability is .26 (39 divided by 150) that a student selected randomly from the sample will report a balance between \$1,500 and \$2,500.

The shape of the probability distribution is the same as the frequency distribution; we have just relabeled the vertical axis. But the probability distribution has two important characteristics not possessed by the frequency histogram: (1) the height of each bar is equal to or greater than 0 and equal to or less than 1 and (2) the sum of the heights of the bars is equal to 1. These characteristics qualify the new display as a

probability distribution for nominal or ordinal variables.²The probability of an outcome is defined as ranging between 0 and 1, and the sum of probabilities across all possible outcomes is 1.

The probability distribution in figure 5.2 is an empirical distribution because it is based on experience. “Theoretical” probability distributions are also important in drawing conclusions from data and deciding actions to take. An example relevant to the decisions that gamblers make is the distribution of possible outcomes from throwing a six-sided die. In theory, the probability distribution for the six possible outcomes could be displayed with six bars, each having a height of 1/6.

Theoretical distributions that play key roles in the methods of inferential statistics are the binomial, normal, chi-square, t, and F distributions. Actually, each of these names refers to a whole family of distributions. The distributions are described in widely available tables that give numerical information about the distributions. For tables and discussions of the distributions, consult a statistics text such as Loether and McTavish (1988). For example, one could use a table of the normal distribution (with mean of 0 and standard deviation of 1) to find the probability that an observation from a population with this distribution could exceed a specified value. Before computers became commonplace for statistical calculations, tables of the distributions were indispensable to the application of inferential statistics.

²Nominal and ordinal variables take on a finite set of values. Interval-ratio variables have a potentially infinite set of values, so the corresponding probability distribution is defined a little differently. (These variables are introduced under “Level of Measurement” in chapter 1.)

Sampling Distributions

The distribution of responses from 150 college students in the example above is the distribution of a sample. If we were to draw another sample of 150 students and plot a histogram, we would almost surely see a slightly different distribution and the mean would be different. And we could go on drawing more samples and plotting more histograms. Differences among the resulting distributions of samples are inherent in the sampling process.

The aim is to be precise about how much variation to expect among statistics computed from different samples. For example, if we use the mean of a sample to describe the distribution of loan balances in a student population, how much uncertainty derives from using a sample? New kinds of distributions called sampling distributions of statistics, or just sampling distributions for short, provide the basis for making statements about statistical uncertainty.

To this point, we have computed statistics without concern for how we produced the data but now we must use probability sampling, which requires that data be produced by a random process. In particular, suppose that we were to draw 100 different simple random samples, each with 150 students, and compute sample statistics, such as the mean, for each sample.³ This would give us a data sheet like that in table 5.1. Since the computed sample means vary across the samples, we could draw a histogram showing the distribution of the sample means (figure 5.3). The midpoint of each bar along the X axis is the midpoint of an interval centered on the number shown. Such a distribution is what we mean by a sampling distribution—one that tells us the probability of obtaining a sample in which a computed statistic, such as the mean, will have

³There are many kinds of probability samples. The most elementary is the simple random sample in which each member of the population has an equal chance of being drawn to the sample.

certain values.⁴ Using figure 5.3, we can say that 25 percent of the sample means had values in the interval \$3,000 plus or minus 50. Using such information, we will be able to make a statement about the probability that a given interval includes the value of the population mean.⁵ This idea is developed further in a later section on interval estimation.

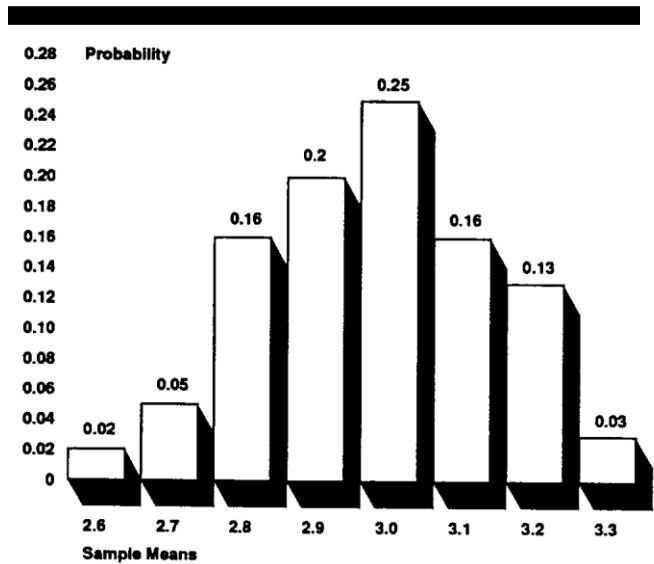
**Table 5.1: Data Sheet for
100 Samples of College
Students**

Sample	Computed mean loan balance
1	\$2,907
2	2,947
3	2,933
4	3,127
5	3,080
100	3,227

⁴Notice the difference between a sample distribution (the distribution of a sample) and a sampling distribution (the distribution of a sample statistic).

⁵The mean either lies in a given interval or it does not. No probability is involved in that respect. However, the probability statement is appropriate since the population mean is usually unknown and we use the confidence interval as a measure of the uncertainty in our estimate of the mean that stems from sampling.

**Figure 5.3: Sampling
Distribution for Mean
Student Loan Balances**



Speaking practically, of course, we would not want to draw many samples of college students because a principal reason for sampling, after all, is to avoid having to make a large number of observations. Therefore, we cannot hope actually to produce a distribution like that of figure 5.3 from empirical evidence. But if our sample is a probability sample, we can usually determine the amount of uncertainty associated with sampling and yet draw only one sample. With a probability sample, the laws of probability often enable us to know the theoretical distribution of a sample statistic so that we can use that instead of an empirical distribution obtained by drawing many samples.⁶

⁶This is where families of distributions like the chi-square and the t come into play to help us estimate population parameters. They are the theoretical distributions that we need.

The sample displayed in figure 5.1, as well as the other 99, was, in fact, drawn randomly from a population with a mean loan balance of \$3,000. It can therefore be used to estimate population parameters for the distribution of students.

Population Parameters

A population parameter is a number that describes a population. Consider again the question of the mean student loan balance for college students. We want to know about the population of all college students—specifically, we want to know the mean loan balance—but we do not want to get information from all. We describe the situation by saying that we want to estimate a population parameter—in this case, the mean of the distribution of loan balances for all students. We want a reasonably close estimate but we are willing to tolerate some uncertainty in exchange for avoiding the cost and time of querying every college student.

The idea of a population parameter applies to any variable measured on a population and any single number that might be used to describe the distribution of the variable. For example, if we want to estimate the proportion of households that use food stamps among those eligible to receive them, the population is all the eligible households. The response variable, use of food stamps, is measured at the nominal level and can have only two values: no or yes. (For purposes of statistical analysis, the variable can be coded as no = 0 and yes = 1.) The population parameter in question is the proportion of all eligible households that use food stamps. The proportion of food stamp users is a way of describing the population so it qualifies as a population parameter.

A population might also be described by two or more variables. For example, we might wish to describe the

population of U.S. households by “use of Medicaid benefits” and “size of the household.” We can deal with the two variables individually, estimating the proportion of households that receive Medicaid benefits on the one hand and estimating the median household size on the other, but we can also estimate a measure of association between two variables.

Parameters are associated with populations, and statistics are associated with samples, but the two concepts are linked in that statistics are used to estimate parameters. Two kinds of estimates for population parameters are possible: point estimates and interval estimates. Both kinds of estimates are statistics computed from probability samples. In the following sections, we first give examples of parameter estimates and then discuss what they mean and how they are computed from samples.

Point Estimates of Population Parameters

A point estimate is a statistic, our “best” judgment about the value of the population parameter in question. In the student loan example, we would like to know the mean loan balance for all students. We draw a simple random sample and use the mean of the sample, a statistic, to estimate the unknown population mean. The value of the sample mean, \$2,907, from the first sample of students is a point estimate of the mean of the population.

The statistical practice is that the sample mean is used to estimate the population mean when a simple random sample is used to produce the data. The procedure has intuitive appeal because the sample mean is the analogue to the population mean. That is, the population mean would be the arithmetic average of all members of the population while the sample

mean is the arithmetic average of all members of the sample.⁷

Point estimates are not based on intuition, however. When a sample has been produced by a random process, statistical theory gives us a good way to estimate a population parameter (that is, theory gives us an appropriate sample statistic for estimating the parameter). That is one of the advantages of randomness; by means of statistical theory, the process provides us with a way to make point estimates of parameters. And it should be noted that intuition does not always suggest the best statistic. For example, intuition might say to estimate the standard deviation of a normally distributed population with a sample standard deviation. However, theory tells us that with small samples, the sample standard deviation should not be used to estimate the standard deviation of the population.

Like the mean and standard deviation, other population parameters are estimated from sample statistics. For example, to answer the question about the proportion of households that are eligible for food stamps, we could use the proportion eligible from a simple random sample of households to make a point estimate of the proportion eligible in the population. Study questions might require that we estimate a variety of population parameters, including the spread of a distribution and the association between two variables.

One of the important factors determining the choice of a statistic to estimate a population parameter is the

⁷Note that the use of the sample mean to estimate the population mean does not deal with the question, raised in chapter 2, as to the circumstances under which the mean is the best measure of central tendency. When the population distribution is highly asymmetric, the population median may be a better measure of central tendency for some purposes. We would then want point and interval estimates of the median.

procedure used to produce the data—that is, the sample design. To use the methods of statistical inference, the sampling procedure must involve a random process but that still leaves many options (see the transfer paper entitled Using Statistical Sampling). In the student loan example, the sample design was a simple random sample, and that allowed the use of the sample mean to estimate the population mean. With the simple random sample, each student in the population had an equal probability of being drawn to the sample, but as a practical matter such samples are not often used. Instead, commonly used sample designs such as a stratified random sample imply unequal, but known, probabilities so that a weighted sample mean is used to estimate the population mean. The procedures for estimating the population mean then become a little more complicated (for example, we have to determine what weights to use) but the statistical principles are the same.

A point estimate provides a single number with which to describe the distribution of a population. But as we have seen in table 5.1, different samples yield different numerical values that are not likely to correspond exactly to the population mean. We sample because we are willing to trade off a little error in the estimate of the population parameter in exchange for lower cost. But how much error should we expect from our sampling procedure? Interval estimates, the subject of the next section, enable us to describe the level of sampling variability in our procedures.

GAO reports provide numerous illustrations of point estimates of population parameters. The most commonly estimated parameters are probably the mean of a normal distribution and the probability of an event in a binomial distribution. In a study of the

Food Stamp program, for example, the probability of program participation was estimated for all eligible households and many subcategories of households (U.S. General Accounting Office, 1990b). The estimates were based upon a nationally representative sample of 7,061 households. A study of bail reform estimated the mean number of days in custody for two groups of felony defendants in four judicial districts (U.S. General Accounting Office, 1989b). Population means for two 6-month periods in 1984 and 1986 were estimated from two probability samples of 605 and 613 defendants, respectively.

Interval Estimates of Population Parameters

Point estimates of population parameters are commonly made and, indeed, sometimes only point estimates are made. That is unfortunate because point estimates are apt to convey an unwarranted sense of precision. A point estimate should be accompanied by interval estimates to show the amount of variability in the point estimate.

An interval estimate of a population parameter is composed of two numbers, called lower and upper confidence limits, each of which is a statistic. For example, an interval estimate of mean student loan balance is \$2,625 and \$3,189, corresponding to the two limits. Formulas for computing confidence limits are known for many population parameters (see statistical texts such as Loether and McTavish, 1988).

To interpret an interval estimate properly, we need to imagine drawing multiple samples. Following our student loan example, we can suppose that we have 100 samples and construct an expanded version of the data sheet in table 5.1. The interval based on the first sample is in row 1 of table 5.2 and we have computed intervals for each of the 5 other samples in the display. If the table were completely filled out, we

would have estimates for 100 intervals just as we have 100 point estimates.

Table 5.2: Point and Interval Estimates for a Set of Samples

Sample	Point estimate	Interval estimate
1	\$2,907	\$2,625-3,189
2	2,947	2,667-3,227
3	2,933	2,647-3,219
4	3,127	2,947-3,397
5	3,080	2,810-3,350
100	3,227	2,959-3,595

An interval estimate has the following interpretation: among all the interval estimates made from many samples of a population, approximately P percent will enclose the true value of the population parameter. The value of P is the confidence level and is frequently set at .95. With respect to the interval estimates in table 5.2, this means that approximately 95 out of 100 intervals are boundaries of the true value of the population parameter.

Because we do not actually draw 100 samples, we must now translate the foregoing reasoning to the situation in which we draw a single sample. Suppose it is sample 1 in table 5.2. This sample produced lower and upper bounds of \$2,625 and \$3,189. Following the reasoning above and assuming this is the only sample drawn, we would say that we are 95-percent confident that the mean loan balance is between \$2,625 and \$3,189. That is, applying the interval-estimating procedure to all possible samples, a statement that a given interval enclosed the mean would be correct 95 percent of the time. Therefore, for our single sample, we are justified in claiming that we are 95-percent sure that it embraces the true population mean. We must always admit that if we are unlucky,

our estimate based upon sample 1 might be one of the 5 percent that does not bound the population mean.

To make an interval estimate, we choose a confidence level and then use the value to calculate the confidence limits. P can be any percentage level, up to almost 100, but by convention it is usually set at 90 or 95. The larger the value of P , the wider will be the interval estimate. In other words, to increase the likelihood that an interval will “cover” the population parameter, the interval must be widened.

The interval estimate has intuitive appeal because when the confidence level is high, say 95 percent, we feel that the population parameter is somewhere within the interval—even though we know that it might not be.

As in our discussion of sampling distributions, the idea of drawing multiple samples is only to further our understanding of the underlying principle. To actually make an interval estimate, we draw one sample and use knowledge of probability and theoretical sampling distributions to compute the confidence limits. For example, we know from the central limit theorem of probability theory that if the sample size is relatively large (say greater than 30), then the sampling distribution of sample means is distributed approximately as a normal distribution, even if the distribution of the population is not.⁸ Then we can use formulas from probability theory and published tables for the t distribution to compute the lower and upper confidence limits. Of course, in practice the calculations are usually carried out on a computer that is simply given instructions to carry out all or most of the steps necessary to produce an interval estimate from the sample data. It should be

⁸Notice that although the distribution of loan balances in figure 5.2 is somewhat asymmetric, the sampling distribution is more symmetric.

noted, however, that the computer does not know whether the data were produced by a random process. It will analyze any set of data; the analyst is responsible for ensuring that the assumptions of the methodology are satisfied.

Statistical analyses similar to the one just outlined for estimating a population mean can be used to estimate other parameters such as the spread in the amount of hazardous waste produced by generators or the association between the use of Medicaid benefits in a household and the size of the household. For the hazardous waste question, we might obtain an interval estimate of the standard deviation (see “Measures of the Spread of a Distribution” in chapter 3), and for the Medicaid question we probably would make an interval estimate for the point biserial correlation (see “The Comparison of Groups” in chapter 4).⁹

An interval estimate allows us to express the uncertainty we have in the value of a population parameter because of the sampling process but it is important to remember that there are other sources of uncertainty. For example, measurement error may substantially broaden the band of uncertainty regarding the value of a parameter.

The GAO studies cited earlier as illustrating point estimates also provide examples of interval estimates. Confidence intervals were estimated for the probability of Food Stamp program participation (U.S. General Accounting Office, 1990b) and for the mean days spent in custody by felon defendants (U.S. General Accounting Office, 1989b).

⁹Obtaining an interval estimate for the standard deviation is highly problematic because, unlike the case of the mean, the usual procedures are invalid when the distribution of the variable is not normal.

Determining Causation

“Correlation does not imply causation” is a commonly heard cautionary statement about a correlation or, more generally, an association between two variables. But causation does imply association. That is, if two variables are causally connected, they must be associated—but that is not enough. In this chapter, we consider the evidence that is necessary to answer questions about causation and, briefly, some analytical methods that can be brought to bear. In other words, we address the fourth and final generic question in table 1.3.

The following example, similar to one given in chapter 4, is a question framed in causal terms:

- Are homeowners’ appliance-purchasing decisions affected by government information campaigns aimed at reducing energy consumption? (Note the substitution of “affected by” for “associated with.”)

Some related questions can be imagined:

- Are homeowners’ attitudes about energy conservation influenced by their income level?
- Do homeowners purchase energy-efficient appliances as a consequence of government-required efficiency labels?
- Is the purchase of energy-efficient appliances causally determined by homeowners’ income level?

If it is possible to collect quantitative information on such issues, statistical analysis may play a role in drawing conclusions about causal connections.

What Do We Mean by Causal Association?

What does it mean to say that homeowners' decisions to purchase energy-efficient appliances are affected by a government information campaign? It means that the campaign in some sense determines whether the homeowners purchase energy-efficient appliances.¹ There is thus a link between the campaign and the purchase decisions. To claim a causal link is to claim that exposure to the campaign influences the likelihood that homeowners will purchase energy-efficient appliances. Three aspects of causal links have a bearing on how we analyze the data and how we interpret the results.

First, an association between two variables is regarded as a probabilistic one. For most of GAO's work, associations are not certain. For example, most people exposed to the government energy information campaign might purchase energy-efficient appliances but some might not. So knowing the attribute for one variable does not allow us to predict the attribute of the other variable with certainty. In this paper, we assume that the cause-and-effect variables are expressed numerically with the consequence that statistical methods can be used to analyze probabilistic associations. In particular, measures of association indicate the strength of causal connections.

Second, a causal association is temporally ordered. That is, the cause must precede the effect in time. Perhaps income causes attitude about conservation or, conceivably, attitude causes income—but it does not work both ways at exactly the same time.² This

¹The exact nature of causation, both physical and social, is much debated. We do not delve into the intricacies in this paper. There are many detailed discussions of the issues; Bunge (1979) and Hage and Meeker (1988) are two.

²The asymmetry feature does not rule out reciprocal effects in the sense that first attitude affects income, then income affects attitude, and so on.

means that, for causation to be established, the relationship between two variables must be asymmetric, whereas a measure of association between two variables can be either symmetric or asymmetric. In statistical language, the direction of causality is expressed by referring to the cause variable as the independent variable and the effect variable as the dependent variable. Measures like those in chapter 4, if they are asymmetric ones, are used to characterize the association between dependent and independent variables.

Third, we must assume that an effect has more than one cause or that a cause has more than one effect. In the real world, a causal process is seldom if ever limited to two variables. It seems likely that a number of factors would influence a homeowner's purchasing decisions—knowledge acquired from the government information program perhaps, but also maybe income and educational level. It is also likely that the decision would vary by the homeowners' age, gender, place of residence, and probably many more factors. In trying to determine the extent of causal association between any two variables, we have to consider a whole network of associations. If we look only at the association between exposure to the government program and the purchase decision, we are likely to draw the wrong conclusion.

Evidence for Causation

Thus, determining the causal connection between two variables is a formidable task involving a search for evidence on three conditions: (1) the association between two variables, (2) the time precedence between them, and (3) the extent to which they have been analyzed in isolation from other influential

variables.³ In short, analyzing causation requires evidence on the association and time precedence of isolated variables.⁴

When we speak of evidence about the association between two variables, we mean simply that we can show the extent to which a variable X is associated with another variable Y. Asymmetric measures of association provide the necessary evidence. If we treat X as the independent variable and Y as the dependent one, compute an appropriate measure, and find that it is sufficiently different from zero, we will have evidence of a possibly causal relation.⁵ For example, if we had data on whether homeowners were exposed to a government program that provided energy information and the extent to which they have purchased energy-efficient appliances, we could compute a measure of association between the two variables. However, a simple association between two variables is usually not sufficient, because other variables are likely to influence the dependent variable, and unless we take them into account, our

³The three conditions are almost uniformly presented as those required to “establish” causality. However, the language varies from authority to authority. This paper follows Bollen (1989) in using the concept of isolation rather than that of nonspuriousness, the more usually employed concept.

⁴In this chapter, we discuss evidence for a causal relationship between quantitative variables and methods, as used in program evaluation and the sciences generally, for identifying causes. The word “cause” is used here in a more specific way than it is used in auditing. There, “cause” is one of the four elements of a finding, and the argument for a causal interpretation rests essentially on plausibility rather than on establishing time-ordered association and isolating a single cause from other potential ones. The methods described in this paper may help auditors go beyond plausibility arguments in the search for causal explanations. See U.S. General Accounting Office, *Government Auditing Standards* (Washington, D.C.: 1988), standard 11 on page 6-3 and standards 21-24 on page 7-5.

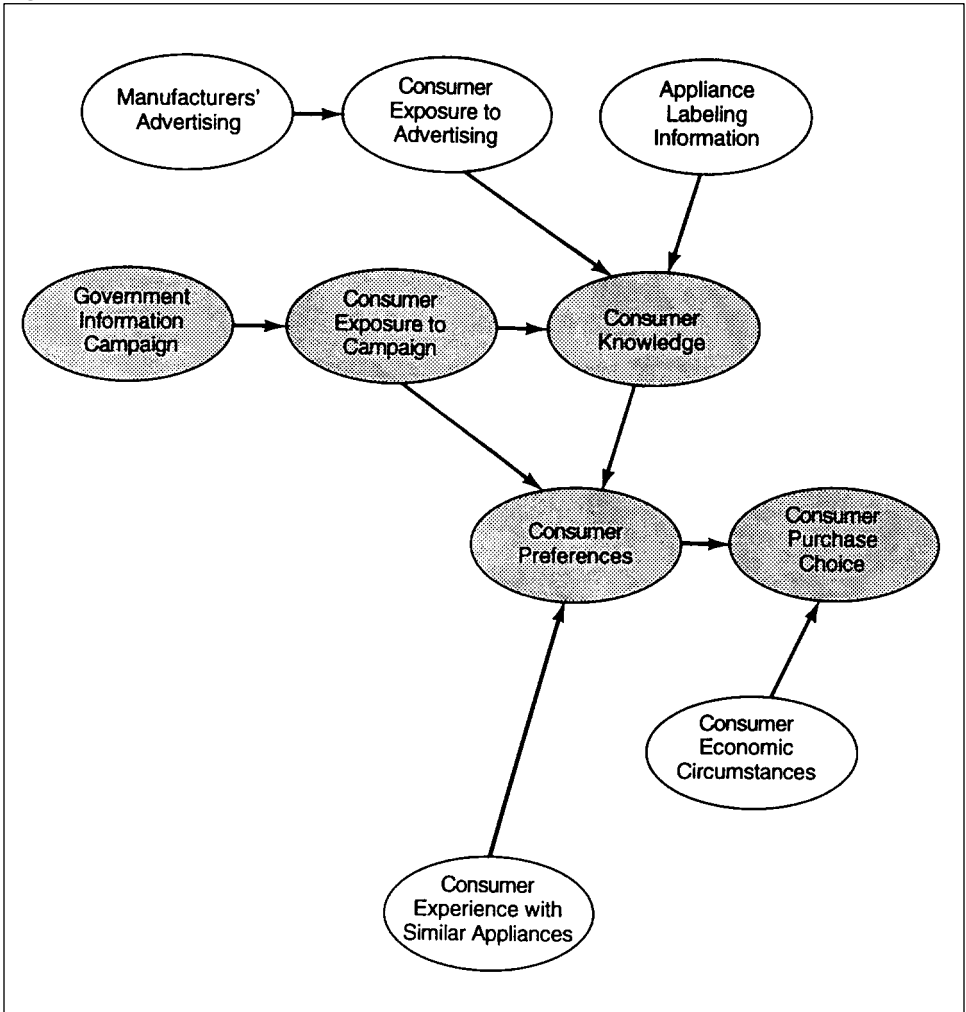
⁵Judgment is applied in deciding the magnitude of a “sufficient difference.”

estimate of the extent of the causal association will be wrong.

Taking account of the other variables means determining the association between X and Y in isolation from them. This is necessary because in the real world, as we have noted, the two variables of interest are ordinarily part of a causal network—with perhaps many associated variables including several causal links.

Figure 6.1 shows a relatively small network of which our two variables, consumer-exposure-to-campaign and consumer-purchase-choice, are a part. The arrows in the network indicate possible causal links. The government information campaign plus variables that may be affected by it are indicated by shaded areas. Other variables that may influence the consumer's choice of appliance are represented by unshaded areas.

Figure 6.1: Causal Network



Consumer-exposure-to-campaign and consumer-purchase-choice may indeed have an underlying causal association, but the presence of the other variables will distort the computed association unless we isolate the variables. That is, the computed amount of the association between X and Y may be either greater or less than the true level of association unless we take steps to control the influence of the other variables. Control is exerted in two ways: by the design of the study and by the statistical analysis.

Finally, we must also have evidence for time precedence, which means that we must show that X precedes Y in time. If we can show that the appliance purchases always came after exposure to the information program, then we have evidence that X preceded Y. Note that the use of asymmetric measures of association does not ensure time precedence. We can compute asymmetric measures for any pair of variables. Evidence for time precedence comes not from the statistical analysis but, rather, from what we know about how the data on X and Y were generated.

Determining the association between two variables is usually not much of a technical problem because computer programs are readily available that can calculate many different measures of association. Establishing time precedence can sometimes be difficult, depending in part upon the type of design employed for the study. (See the transfer paper entitled Designing Evaluations. For example, with a cross-sectional survey, it may not be easy to decide which came first—a consumer's preference for certain appliances or exposure to a government information program. But with other designs, like an experiment that exposes people to information and then measures their preference, the evidence for time precedence may be straightforward.

Most difficulties in answering causal questions (sometimes called “impact” questions) stem from the requirement to isolate the variables. In fact, it is never possible to totally isolate two variables from all other possible influences, so it is not possible to be absolutely certain about a causal association. Instead, the confidence that we can have in answering a causal question is a matter of degree—depending especially upon the design of the study and the kind of data analysis conducted.

The key task of isolating two variables—or, in other words, controlling variables that confound the association we are interested in—can be approached in a variety of ways. Most important is the design for producing the data. For simplicity, consider just two broad approaches: experimental and nonexperimental designs.

In the most common type of experiment, we form two groups of subjects or objects and expose one group to a purported cause while the other group is not so exposed. For example, one group of homeowners would be exposed to a government information campaign about energy conservation and another group would not be exposed. In data analysis terms, we would thus have a nominal, independent variable (X), usually called a treatment, that has two attributes: exposure-to-the-campaign and nonexposure-to-the-campaign. If the groups are formed by random assignment, the design is called a true experiment; otherwise, it is called a quasi-experiment.

In answering a causal question based upon experimental data, our basic logic is to compare what happens to the dependent variable Y when the purported cause is present ($X = 1$) with what happens when it is absent ($X = 0$). For example, we could

compare the overall proportion of energy-efficient appliances purchased by the two groups. In a true experiment, isolation is achieved by the process of random assignment, which ensures that the two groups are approximately equivalent with respect to all variables, except X, that might affect the purchase of appliances. In this sense, the variables Y and X have been isolated from the other variables and a measure of association between Y and X can be taken as a defensible indicator of cause and effect.

In a quasi-experiment, random assignment is not used to form the two groups but, rather, they are formed or chosen so that the two groups are as similar as possible. The quasi-experimental procedure, while imperfect, can isolate X and Y to a degree and may provide the basis for estimating the extent of causal association.

In a nonexperimental design, there is no effort to manipulate the purported cause, as in a true experiment, or to contrive a way to compare similar groups, as with a quasi-experiment. Observations are simply made on a collection of subjects or objects with the expectation that the individuals will show variation on the independent and dependent variables of interest. Sample surveys and multiple case studies are examples of nonexperimental designs that could be used to produce data for causal analysis.⁶ For example, we might conduct telephone interviews with a nationally representative sample of adults to learn about their attitudes toward energy conservation and the extent to which they are aware of campaigns to reduce energy use. The designs for sample surveys and case studies do not isolate the key variables, so

⁶Sample surveys and case studies can be used in conjunction with experimental designs. For example, a sample survey could be used to collect data from the population of people participating in an experiment.

the burden falls on the data analysis, a heavy burden indeed.

Two broad strategies for generating evidence on the association and time precedence of isolated variables are available: experimental and nonexperimental. Using such evidence, data analysis aimed at determining causation can be carried out in a variety of ways. As noted, we are here assuming that the data are quantitative. Other approaches are necessary with qualitative data. (Tesch, 1990, describes computer programs such as AQUAD and NUDIST that have some capability for causal analysis.)

Causal Analysis of Nonexperimental Data

All analysis methods involve determining the time order and the extent of any association between two variables while attempting to isolate those two variables from other factors. While it might be tempting just to compute an asymmetric measure of association between the variables—for example, by determining the regression coefficient of X when Y is regressed on X—such a procedure would almost always produce misleading results. Rather, it is necessary to consider other variables besides X that are likely to affect Y.

The preferred method of analysis is to formulate a causal network—plausible connections between a dependent variable and a set of independent variables—and to test whether the observed data are consistent with the network.⁷ There are many related ways of doing the testing that go by a variety of

⁷Unless the causal network is an unusually simple one, just adding additional variables to the regression equation is not an appropriate form of analysis.

names, but structural equation modeling seems the currently preferred label.⁸

Two distinct steps are involved in structural equation modeling. The first is to put forth a causal network that shows how the variables we believe are involved in a causal process relate to one another (figure 6.1). The network, which can take account of measurement error, should be based on how we suppose the causal process works (for example, how a government program X is intended to bring about a desirable outcome Y). This preliminary understanding of causation is usually drawn from evidence on similar programs and from more general research evidence on human behavior and so on.

The second step is to analyze the data, using a series of linear equations that are written to correspond to the network. Computer programs, such as LISREL and EQS, are then used to compare the data on the observed variables with the model and to produce measures of the extent of causal association among the variables. The computer programs also produce indicators of the degree of “fit” between the model and the data. If the fit is not “good” enough, the causal network may be reformulated (step 1) and the analysis (step 2) carried out again. The analyst may cycle through the process many times.

A good fit between the model and the data implies not that causal associations estimated by structural equation modeling are correct but just that the model is consistent with the data. Other models, yet untested, may do as well or better.

With data generated from nonexperimental designs, the statistical analysis is used in an effort to isolate the variables. With experimental designs, an effort is

⁸Other common names for the methods are “analysis of covariance structures” and “causal modeling.”

made to collect the data in such a way as to isolate the variables.

Causal Analysis of Experimental Data

As noted earlier, in a true experiment, random assignment of subjects or objects to treatment and comparison groups provides a usually successful way to isolate the variables of interest and, thus, to produce good answers to causal questions.⁹ In a quasi-experiment, the comparison group is not equivalent (in the random assignment sense) to the treatment group, but if it is similar enough, reasonably good answers to causal questions may be obtained.

The usual ways to analyze experimental data are with techniques such as analysis of variance (ANOVA), analysis of covariance (ANCOVA), and regression. Regression subsumes the first two methods and can be used when the dependent variable is measured at the interval level and with independent variables at any measurement level. If the dependent variable is

⁹An experiment ordinarily provides strong evidence about causal associations because the process of random assignment ensures that the members of treatment and control groups are approximately equivalent with respect to supplementary variables that might have an effect on the response variable. Being essentially equivalent, almost all variables except the treatment are neutralized in that treatment and control group members are equally affected by those other variables. For example, even though a variable like a person's age might affect a response variable such as health status, random assignment would ensure that the treatment and control groups are roughly equivalent, on the average, with respect to age. In estimating the effect of a health program, then, the evaluator would not mistake the effect of age on health condition for a program effect.

measured at the nominal or ordinal level, other techniques such as logit regression are required.¹⁰

Although the experimental design is used in an effort to isolate the variables, the objective is never perfectly achieved. Quasi-experimental designs, especially, may admit alternative causal explanations. Therefore, structural equation modeling is sometimes used to analyze experimental data to further control the variables.

Limitations of Causal Analysis

Statistics texts cover the many assumptions and limitations associated with quantitative analysis to determine causation. The bibliography lists several that give detailed treatments of the methods. However, two more general points need to be made.

First, some effects may be attenuated or changed because of the settings in which they occur—that is, whether the causal process happens in a natural way or is “forced.” In a natural setting, X may have a strong causal influence on Y, but if the setting is artificial, the connection may be different. For example, homeowners who are provided information indicating the advantages of conserving energy (X) may decline to take energy-saving steps (Y) if they are part of a designed experiment. However, the same homeowners might adopt conservation practices if they sought out the information on their own. Strictly speaking, the nature of the X variable is different in these two situations but the point is still the same: the causal process may be affected by differences, sometimes subtle, between the experimental and natural conditions. For some variables, the causal link

¹⁰The line between ordinal and interval data is not hard and fast. For example, many analysts with a dependent variable measured at the ordinal level use regression analysis if they believe the underlying variable is at the interval level (and limited only to ordinal because of the measuring instrument).

might be strongest in the natural setting, but for other variables it might be strongest in the experimental setting.

The second point is that causal processes may not be reversible or they may revert to an original state slowly. To illustrate, suppose that laws to lower the legal age for drinking alcohol have been shown to cause a higher rate of automobile accidents. It does not necessarily follow that subsequent laws to raise the drinking age will produce lower accident rates. Evidence to show the effect of increasing (decreasing) a variable cannot, in general, be used to support a claim about the effects of decreasing (increasing) the variable.

Avoiding Pitfalls

Basic ideas about data analysis have been presented in the preceding chapters. Several methods for analyzing central tendency, spread, association, inference from sample to population, and causality have been broadly described. In keeping with the approach in the rest of the paper, this chapter offers advice at a general level with the understanding that specific strategies and cautions are associated with particular methods.

Attention to data analysis should begin while evaluators are formulating the study questions, and in many instances it should continue until they have made the last revisions in the report. Throughout this time, they have many opportunities to enhance the analysis or to make a misstep that will weaken the soundness of the conclusions that may be drawn.

Analysis methods are intertwined with data collection techniques and sampling procedures so that decisions about data analysis cannot be made in isolation. During the planning stages of a study, evaluators must deal with all three of these dimensions simultaneously; after samples have been drawn and data collected, analysis methods are constrained by what has already happened. If it were necessary to summarize advice in a single word, it would be: anticipate.

In the Early Planning Stages

Be clear about the question. As a study question is being formulated and refined, it helps to think through the implications for data analysis. If evaluators cannot deduce data analysis methods from the question or if the question is so vague as to lead to a variety of possible approaches, then they probably need to restate the question or add some additional statements to elaborate upon the question. For example, a question might be: To what extent have

the objectives of the dislocated worker program been achieved?

By one reading of this question, the appropriate analysis would simply be to determine the extent to which dislocated workers have found new employment at a rate in excess of (or less than) program goals. With proper sampling and data collection, the analysis would be a matter of computing the proportion of a pool of workers who found reemployment and compare that proportion to the goal for the program. This analysis would not permit the policymaker to draw conclusions about whether the program contributed to the achievement of the goal, because the influence of other factors that might affect the reemployment rate have not been considered.

By another reading, the question implies making a causal link between the government program and the proportion of displaced workers who find reemployment. This means that the design and the analysis must contend with the three conditions for causality discussed in chapter 6. For example, an effort must now be made to isolate the two variables, the program and the reemployment rate, from other variables that might have a causal connection with the reemployment rate. The two interpretations of the question are quite different, and so the question must be clarified before work proceeds.

Understand the subject matter. Evaluators usually need in-depth knowledge of the subject matter to avoid drawing the wrong conclusion from a data set. Numbers carry no meaning except that which derives from how the variables were defined. Moreover, data are collected in a social environment that is probably changing over time. Consequently, there is often an

interplay between the subject matter and data analysis methods.

An example using medical data illustrates the importance of understanding the phenomena behind the numbers. Mortality rates for breast cancer among younger women show some decline over time. However, it would be wrong to draw conclusions about the efficacy of treatment on this evidence alone. It is necessary to understand the details of the process that is producing the numbers. One important consideration is that diagnostic techniques have improved so that cancers are detected at an earlier stage of development. As a consequence, mortality rates will show a decline even if treatment has not improved. A data analysis aimed at determining change in mortality from changes in treatment must adjust for the “statistical artifact” of earlier detection. (For an elaboration of this example, see U.S. General Accounting Office, 1987a.)

The need to understand the subject matter implies a thorough literature review and consultation with diverse experts. It may also mean collecting supplementary data, the need for which was not evident at the outset of the study. For example, in a study of cancer mortality rates, it would be necessary to acquire information about the onset of new diagnostic procedures.

Develop an analysis plan. The planning stage of a project should yield a set of questions to be answered and a design for producing the answers. A plan for analysis of the data should be a part of the design.

Yin (1989) has observed that research designs deal with logical problems rather than logistical problems. So it is with the analysis plan—it should carry forward the overall logic of the study so that the

connection between the data that will be collected and the answers to the study questions will become evident. For example, if a sample survey will be used to produce the data, the analysis plan should explain the population parameters to be estimated, the analysis methods, and the form of reporting. Or, if a field experiment will produce the data, the plan should explain the comparisons to be made, the analysis methods to be used, any statistical adjustments that will be made if the comparison groups are nonequivalent, and the form of reporting that will be used. Another matter that should be considered at this time is the appropriate units of analysis. Whatever the nature of the study, the analysis plan should close the logical loop by showing how the study questions will be answered.

When Plans Are Being Made for Data Collection

Coordinate analysis plans with methods for selecting sources of information. The methods for selecting data sources strongly determine the kinds of analysis that can be applied to the resulting data. As noted in earlier chapters, evaluators can use descriptive statistics in many situations, but inferential techniques depend upon knowledge of sampling distributions, knowledge that can be applied only when the data have been produced by a random process.

Random processes can be invoked in many ways and with attendant variations in analytic methods. Often the choice of sampling procedure can affect the efficiency of the study as well. Evaluators should make a decision on the particular form of random selection in consultation with a sampling statistician in advance of data collection. Unless proper records of the sampling process are maintained, an analyst may not be able to use statistical inference techniques to estimate population parameters.

Coordinate analysis plans with data collection. As data collection methods are firmed up and instruments are developed, variables will be defined and measurement levels will be determined. This is the time to review the list of variables to ensure that all those necessary for the analysis are included in the data collection plans. The measurement level corresponding to a concept is often intrinsic to the concept, but if that is not so, it is usually wise to strive for the higher levels of measurement. There may be analytic advantage to the higher levels or, if going to a higher level is more costly, the proper trade-off may be to settle for a somewhat weaker analysis method.¹

As the Data Analysis Begins

Check the data for errors and missing attributes. No matter how carefully evaluators have collected, recorded, and transformed the data to an analysis medium, there will be errors. They can detect and remove some by simple checks. Computer programs can be written, or may already exist, for checking the plausibility of attributes. For example, the gender variable has two attributes, male and female, and therefore two possible numerical values, say 0 and 1. Any other value is an error and can be readily detected. In a similar way, evaluators can check all variables to ensure that the attributes in the data base are reasonable.

They can detect other errors by contingency checks. Such checks are based on the fact that the attributes for some variables are conditional upon the attributes

¹Flexibility usually exists on the fuzzy border between ordinal and interval variables. Analysts often treat an ordinal variable as if it were measured at the interval level. In fact, some authorities (see Kerlinger, 1986, pp. 401-3, for example) believe that most psychological and educational variables approximate interval equality fairly well. In any case, instrument construction should take account of the measurement level desired.

of other variables. For example, if a medical case has “male” as an attribute for gender, then it should not have “pregnant” as an attribute of physical condition. These kinds of if-then checks on the attributes are also relatively easy to automate.

A missing attribute, where none of the acceptable attributes for the variable is present, is more difficult to deal with. Evaluators have four options: (1) go back to the data source and try to recover the missing attribute, (2) drop the case from all analyses, (3) drop the case from any analysis involving the variable in question but use the case for all other analyses, and (4) fill in a substitute value for the missing attribute. Considerations involved in dealing with missing attributes are treated by many writers. (See, for example, Groves, 1989; Little and Rubin, 1987; and Rubin, 1987.)

When evaluators have used probability sampling with the aim of estimating population parameters from sample results, overall nonresponse by units from the sample is an especially important problem. If the nonresponse rate is substantial and if it cannot be shown that the respondents and nonrespondents are probably similar on variables of interest, doubt is cast on the estimates of population parameters. Consequently, an analysis of nonrespondents will be needed. See Groves (1989) for an introduction to the literature on nonresponse issues.

Explore the data. A number of statistical methods have been specifically developed to help evaluators get a feel for the data and to produce statistics that are relatively insensitive to idiosyncracies in the data. Some of these, like the stem-and-leaf plot and the box-and-whiskers plot, are graphic and especially useful in understanding the nature of the data. (Details about exploratory data analysis may be found

in Tukey, 1977; Hoaglin, Mosteller, and Tukey, 1983, 1985; Velleman and Hoaglin, 1981; and Hartwig and Dearing, 1979.)

Fit the analysis methods to the study question and the data in hand. The appropriateness of an analysis method depends upon a number of factors such as the way in which the data sources were selected, the measurement level of the variables involved, the distribution of the variables, the time order of the variables, and whether the intent is to generalize from the cases for which data are available to a larger population. Some factors, like the measurement level, must be considered in every data analysis, while others, like time order, may be germane only for certain types of questions—in this instance, a question about a causal association.

When evaluators consider two or more different analysis methods, the choice may not be obvious. For example, with interval level measurement, the median may be preferable to the mean as a measure of central tendency if the distribution is very asymmetrical. But asymmetry is a matter of degree and a little error from asymmetry may be acceptable if there are strong advantages to using the mean. Or it may be easy to transform the variable so that near symmetry is attained. Statistical tests that indicate the degree of asymmetry are available, but ultimately the evaluators have to make a judgment.

“The data don’t remember where they came from.” These words of a prominent statistician underscore the point that the data analyst must be mindful of the process that generated the data. We can blindly apply a host of numerical procedures to a data set but many of them would probably not be appropriate in view of the process that produced the data. For example, the methods of statistical inference apply only to data

generated by a random process or one that is “random in effect.” (For a discussion of the circumstances in which statistical inference is appropriate, see Mohr, 1990, pp. 67-74.) Since the data do not remember how they were produced, the analyst has to ensure that the techniques are not misapplied.

Monitor the intermediate results and make analytic adjustments as necessary. Even with good planning, it is not possible to foresee every eventuality. The data in hand may be different from what was planned, or preliminary analyses may suggest new questions to explore. For example, the distribution of the data may take a form not anticipated so that analytic transformations are necessary. Or, a program may have an unanticipated effect that warrants a search for an explanation. The analyst must scrutinize the intermediate results carefully to spot opportunities for supplementary analyses as well as to avoid statistical procedures that are not compatible with the data.

As the Results Are Produced and Interpreted

Use graphics but avoid displays that distort the data. The results of quantitative data analysis may be terse to the point of obtuseness. Graphics may help both in understanding the results and in communicating them. There are many excellent examples of how to visually display quantitative information but even more of how to distort and obfuscate. (For introductions to graphic analysis and data presentation, see Cleveland, 1985; Du Toit, Steyn, and Stumpf, 1986; and Tufte, 1983.)

Be realistic and forthright about uncertainty. Uncertainty is inherent in real-world data. All measurements have some degree of error. If sampling is used, additional error is introduced. Data entry and data processing may produce yet more error. While

evaluators can and should take steps to reduce error, subject to resource constraints, some error will always remain. The question that must be addressed is whether the level of error present threatens what are otherwise the conclusions from the study.

A complementary question is how to report the nature and extent of error. Reporting issues for some forms of quantitative analysis have been given considerable attention and several professional organizations offer guidelines.² The basic rule is to be forthright about the nature of the evidence.

²The Evaluation Research Society (now merged with Evaluation Network to become the American Evaluation Association) published standards that include coverage of reporting issues (Rossi, 1982). Other standards that give somewhat more attention to statistical issues are those of the American Association of Public Opinion Research (1991) and the Council of American Survey Research Organizations (1986). In 1988, the federal government solicited comments on a draft Office of Management and Budget circular establishing guidelines for federal statistical activities. A final version of the governmentwide guidelines, which included directions for the documentation and presentation of the results of statistical surveys and other studies, has not been published.

Bibliography

The statistics texts marked with an asterisk (*) provide more detailed information on the statistical topics mentioned in this paper.

American Association of Public Opinion Research. Code of Professional Ethics and Practices. Ann Arbor, Mich.: 1991.

Arney, W. R. Understanding Statistics in the Social Sciences. New York: W. H. Freeman, 1990.*

Bollen, K. A. Structural Equations With Latent Variables. New York: John Wiley, 1989.

Bornstedt, G. W., and D. Knoke. Statistics for Social Data Analysis. Itasca, Ill.: F. E. Peacock, 1982.

Bryman, A., and D. Cramer. Quantitative Data Analysis for Social Scientists. New York: Routledge, 1990.*

Bunge, M. Causality and Modern Science, 3rd ed. rev. New York: Dover Publications, 1979.

Cleveland, W. S. The Elements of Graphing Data. Monterey, Calif.: Wadsworth, 1985.

Cohen, J., and P. Cohen. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1975.*

Council of American Survey Research Organizations. Code of Standards. Port Jefferson, N.Y.: 1986.

Draper, N., and H. Smith. Applied Regression Analysis, 2nd ed. New York: John Wiley, 1981.*

Bibliography

Du Toit, S. H. C., A. G. W. Steyn, and R. H. Stumpf. Graphical Exploratory Data Analysis. New York: Springer-Verlag, 1986.

Freedman, D., R. Pisani, and R. Purves. Statistics. New York: W. W. Norton, 1980.*

Groves, R. M. Survey Errors and Survey Costs. New York: John Wiley, 1989.

Hage, J., and B. F. Meeker. Social Causality. Boston: Unwin Hyman, 1988.

Hahn, G. J., and W. Q. Meeker. Statistical Intervals: A Guide for Practitioners. New York: John Wiley, 1991.*

Hartwig, F., and B. E. Dearing. Exploratory Data Analysis. Newbury Park, Calif.: Sage, 1979.

Hays, W. L. Statistics, 3rd ed. New York: Holt, Rinehart and Winston, 1981.*

Hildebrand, D. K., J. D. Laing, and H. Rosenthal. Analysis of Ordinal Data. Newbury Park, Calif.: 1977.*

Hoaglin, D. C., F. Mosteller, and J. W. Tukey (eds). Understanding Robust and Exploratory Data Analysis. New York: John Wiley, 1983.

Hoaglin, D. C., F. Mosteller, and J. W. Tukey (eds). Exploring Data Tables, Trends, and Shapes. New York: John Wiley, 1985.

Kerlinger, F. N. Foundations of Behavioral Research, 3rd ed. New York: Holt, Rinehart and Winston, 1986.

Little, R. J. A., and D. B. Rubin. Statistical Analysis with Missing Data. New York: John Wiley, 1987.

Loether, H. J., and D. G. McTavish. Descriptive and Inferential Statistics: An Introduction, 3rd ed. Boston: Allyn and Bacon, 1988.

McPherson, G. Statistics in Scientific Investigation: Its Basis, Application, and Interpretation. New York: Springer-Verlag, 1990.*

Madansky, A. Prescriptions for Working Statisticians. New York: Springer-Verlag, 1988.*

Madow, W. G., I. Olkin, and D. B. Rubin. Incomplete Data in Sample Surveys, vols. 1-3. New York: Academic Press, 1983.

Mohr, L. B. Understanding Significance Testing. Newbury Park, Calif.: Sage, 1990.

Moore, D. S., and G. P. McCabe. Introduction to the Practice of Statistics. New York: W. H. Freeman, 1989.*

Mosteller, F., S. E. Fienberg, and R. E. K. Rourke. Beginning Statistics with Data Analysis. Reading, Mass.: Addison-Wesley, 1983.*

Pedhazur, E. J. Multiple Regression in Behavioral Research, 2nd ed. New York: Holt, Rinehart and Winston, 1982.*

Reynolds, H. T. Analysis of Nominal Data, 2nd ed. Newbury Park, Calif.: Sage, 1984.*

Rossi, P. H. (ed). Standards for Evaluation Practice. San Francisco: Jossey-Bass, 1982.

Rubin, D. B. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley, 1987.

Bibliography

Siegel, S., and N. J. Castellan, Jr. Nonparametric Statistics for the Behavioral Sciences, 2nd ed. New York: McGraw-Hill, 1988.*

Tesch, R. Qualitative Research: Analysis Types and Software Tools. New York: Falmer Press, 1990.

Tufte, E. R. The Visual Display of Quantitative Information. Cheshire, Conn.: Graphics Press, 1983.

Tukey, J. W. Exploratory Data Analysis. Reading, Mass.: Addison-Wesley, 1977.

U.S. General Accounting Office. Potential Effects of a National Mandatory Deposit on Beverage Containers, GAO/PAD-78-19. Washington, D.C.: December 1977.

U.S. General Accounting Office. States' Experience With Beverage Container Deposit Laws Shows Positive Benefits, GAO/PAD-81-8. Washington, D.C.: December 1980.

U.S. General Accounting Office. An Evaluation of the 1981 AFDC Changes: Final Report, GAO/PEMD-85-4. Washington, D.C.: July 1985.

U.S. General Accounting Office. Medical Devices: Early Warning of Problems Is Hampered by Severe Underreporting, GAO/PEMD-87-1. Washington, D.C.: December 1986.

U.S. General Accounting Office. Cancer Patient Survival: What Progress Has Been Made? GAO/PEMD-87-13. Washington, D.C.: March 1987a.

U.S. General Accounting Office. Noncash Benefits: Methodological Review of Experimental Valuation Methods Indicates Many Problems Remain,

Bibliography

GAO/PEMD-87-23. Washington, D.C.:
September 1987b.

U.S. General Accounting Office. The H-2A Program:
Protections for U.S. Farmworkers, GAO/PEMD-89-3.
Washington, D.C.: October 1988.

U.S. General Accounting Office. Children and Youths:
About 68,000 Homeless and 186,000 in Shared
Housing at Any Given Time, GAO/PEMD-89-14.
Washington, D.C.: June 1989a.

U.S. General Accounting Office. Criminal Justice:
Impact of Bail Reform in Selected District Courts,
GAO/GGD-90-7. Washington, D.C.: November 1989b.

U.S. General Accounting Office. Alternative
Agriculture: Federal Incentives and Farmer's
Opinions, GAO/PEMD-90-12. Washington, D.C.:
February 1990a.

U.S. General Accounting Office. Food Stamp
Program: A Demographic Analysis of Participation
and Nonparticipation, GAO/PEMD-90-8. Washington,
D.C.: January 1990b.

U.S. General Accounting Office. Promising Practice:
Private Programs Guaranteeing Student Aid for
Higher Education, GAO/PEMD-90-16. Washington,
D.C.: June 1990c.

U.S. General Accounting Office. Voting: Some
Procedural Changes and Informational Activities
Could Increase Turnout, GAO/PEMD-91-1.
Washington, D.C.: November 1990d.

Velleman, P. F., and D. C. Hoaglin. Applications,
Basics, and Computing of Exploratory Data Analysis.
Boston: Duxbury Press, 1981.

Bibliography

Wallis, W. A., and H. V. Roberts. Statistics: A New Approach. Glencoe, Ill.: Free Press, 1956.*

Yin, R. K. Case Study Research: Design and Methods, rev. ed. Newbury Park, Calif.: Sage, 1989.

Glossary

Analysis of Covariance	A method for analyzing the differences in the means of two or more groups of cases while taking account of variation in one or more interval-ratio variables.
Analysis of Variance	A method for analyzing the differences in the means of two or more groups of cases.
Association	General term for the relationship among variables.
Asymmetric Measure of Association	A measure of association that makes a distinction between independent and dependent variables.
Attribute	A characteristic that describes a person, thing, or event. For example, being female is an attribute of a person.
Batch	A group of cases for which no assumptions are made about how the cases were selected. A batch may be a population, a probability sample, or a nonprobability sample, but the data are analyzed as if the origin of the data is not known.
Bell-Shaped Curve	A distribution with roughly the shape of a bell; often used in reference to the normal distribution but others, such as the t distribution, are also bell-shaped.
Bivariate Data	Information about two variables.
Box-And-Whisker Plot	A graphic way of depicting the shape of a distribution.
Case	A single person, thing, or event for which attributes have been or will be observed.

Glossary

Causal Analysis	A method for analyzing the possible causal associations among a set of variables.
Causal Association	A relationship between two variables in which a change in one brings about a change in the other.
Central Tendency	General term for the midpoint or typical value of a distribution.
Conditional Distribution	The distribution of one or more variables given that one or more other variables have specified values.
Confidence Interval	An estimate of a population parameter that consists of a range of values bounded by statistics called upper and lower confidence limits.
Confidence Level	A number, stated as a percentage, that expresses the degree of certainty associated with an interval estimate of a population parameter.
Confidence Limits	Two statistics that form the upper and lower bounds of a confidence interval.
Continuous Variable	A quantitative variable with an infinite number of attributes.
Correlation	(1) A synonym for association. (2) One of several measures of association (see <u>Pearson Product-Moment Correlation Coefficient</u> and <u>Point Biserial Correlation</u>).

Glossary

Data	Groups of observations; they may be quantitative or qualitative.
Dependent Variable	A variable that may, it is believed, be predicted by or caused by one or more other variables called independent variables.
Descriptive Statistic	A statistic used to describe a set of cases upon which observations were made. Compare with <u>Inferential Statistic</u> .
Discrete Variable	A quantitative variable with a finite number of attributes.
Dispersion	See <u>Spread</u> .
Distribution of a Variable	Variation of characteristics across cases.
Experimental Data	Data produced by an experimental or quasi-experimental design.
Frequency Distribution	A distribution of the count of cases corresponding to the attributes of an observed variable.
Gamma	A measure of association; a statistic used with ordinal variables.
Histogram	A graphic depiction of the distribution of a variable.
Independent Variable	A variable that may, it is believed, predict or cause fluctuation in a dependent variable.

Glossary

Index of Dispersion	A measure of spread; a statistic used especially with nominal variables.
Inferential Statistic	A statistic used to describe a population using information from observations on only a probability sample of cases from the population. Compare with <u>Descriptive Statistic</u> .
Interquartile Range	A measure of spread; a statistic used with ordinal, interval, and ratio variables.
Interval Estimate	General term for an estimate of a population parameter that is a range of numerical values.
Interval Variable	A quantitative variable the attributes of which are ordered and for which the numerical differences between adjacent attributes are interpreted as equal.
Lambda	A measure of association; a statistic used with nominal variables.
Level of Measurement	A classification of quantitative variables based upon the relationship among the attributes that compose a variable.
Marginal Distribution	The distribution of a single variable based upon an underlying distribution of two or more variables.
Mean	A measure of central tendency; a statistic used primarily with interval-ratio variables following symmetrical distributions.

Glossary

Measure	In the context of data analysis, a statistic, as in the expression “a measure of central tendency.”
Median	A measure of central tendency; a statistic used primarily with ordinal variables and asymmetrically distributed interval-ratio variables.
Mode	A measure of central tendency; a statistic used primarily with nominal variables.
Nominal Variable	A quantitative variable the attributes of which have no inherent order.
Nonexperimental Data	Data not produced by an experiment or quasi-experiment; for example, the data may be administrative records or the results of a sample survey.
Nonprobability Sample	A sample not produced by a random process; for example, it may be a sample based upon an evaluator’s judgment about which cases to select.
Normal Distribution (Curve)	A theoretical distribution that is closely approximated by many actual distributions of variables.
Observation	The words or numbers that represent an attribute for a particular case.
Ordinal Variable	A quantitative variable the attributes of which are ordered but for which the numerical differences between adjacent attributes are not necessarily interpreted as equal.

Glossary

Outlier	An extremely large or small observation; applies to ordinal, interval, and ratio variables.
Parameter	A number that describes a population.
Pearson Product-Moment Correlation Coefficient	A measure of association; a statistic used with interval-ratio variables.
Point Biserial Correlation	A measure of association between an interval-ratio variable and a nominal variable with two attributes.
Point Estimate	An estimate of a population parameter that is a single numerical value.
Population	A set of persons, things, or events about which there are questions.
Probability Distribution	A distribution of a variable that expresses the probability that particular attributes or ranges of attributes will be, or have been, observed.
Probability Sample	A group of cases selected from a population by a random process. Every member of the population has a known, nonzero probability of being selected.
Qualitative Data	Data in the form of words.
Quantitative Data	Data in the form of numbers. Includes four levels of measurement: nominal, ordinal, interval, and ratio.

Glossary

Random Process	A procedure for drawing a sample from a population or for assigning a program or treatment to experimental and control conditions such that no purposeful forces influence the selection of cases and that the laws of probability therefore describe the process.
Range	A measure of spread; a statistic used primarily with interval-ratio variables.
Ratio Variable	A quantitative variable the attributes of which are ordered, spaced equally, and with a true zero point.
Regression Analysis	A method for determining the association between a dependent variable and one or more independent variables.
Regression Coefficient	An asymmetric measure of association; a statistic computed as part of a regression analysis.
Resistant Statistic	A statistic that is not much influenced by changes in a few observations.
Response Variable	A variable on which information is collected and in which there is an interest because of its direct policy relevance. For example, in studying policies for retraining displaced workers, employment rate might be the response variable. See Supplementary Variable .
Sample Design	The sampling procedure used to produce any type of sample.

Glossary

Sampling Distribution	The distribution of a statistic.
Scientific Sample	Synonymous with <u>Probability Sample</u> .
Simple Random Sample	A probability sample in which each member of the population has an equal chance of being drawn to the sample.
Spread	General term for the extent of variation among cases.
Standard Deviation	A measure of spread; a statistic used with interval-ratio variables.
Statistic	A number computed from data on one or more variables.
Statistical Sample	Synonymous with <u>Probability Sample</u> .
Stem-And-Leaf Plot	A graphic or numerical display of the distribution of a variable.
Structural Equation Modeling	A method for determining the extent to which data on a set of variables are consistent with hypotheses about causal associations among the variables.
Supplementary Variable	A variable upon which information is collected because of its potential relationship to a response variable.
Symmetric Measure of Association	A measure of association that does not make a distinction between independent and dependent variables.

Glossary

Transformed Variable	A variable for which the attribute values have been systematically changed for the sake of data analysis.
Treatment Variable	In program evaluation, an independent variable of particular interest because it corresponds to a program or a policy instituted with the intent of changing some dependent variable.
Unit of Analysis	The person, thing, or event under study.
Variable	A logical collection of attributes. For example, each possible age of a person is an attribute and the collection of all such attributes is the variable age.

Contributors

Carl Wisler
Lois-ellin Datta
George Silberman
Penny Pickett

Papers in This Series

This is a flexible series continually being added to and updated. The interested reader should inquire about the possibility of additional papers in the series.

The Evaluation Synthesis. Transfer paper 10.1.2.

Content Analysis: A Methodology for Structuring and Analyzing Written Material. Transfer paper 10.1.3.

Designing Evaluations. Transfer paper 10.1.4.

Using Structured Interviewing Techniques. Transfer paper 10.1.5.

Using Statistical Sampling. Transfer paper 10.1.6, formerly methodology transfer paper 6.

Developing and Using Questionnaires. Transfer paper 10.1.7, formerly methodology transfer paper 7.

Case Study Evaluations. Transfer paper 10.1.9.

Prospective Evaluation Methods: The Prospective Evaluation Synthesis. Transfer paper 10.1.10.

Quantitative Data Analysis: An Introduction. Transfer paper 10.1.11.

Ordering Information

The first copy of each GAO report and testimony is free. Additional copies are \$2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. VISA and MasterCard credit cards are accepted, also. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20884-6015

or visit:

Room 1100
700 4th St. NW (corner of 4th & G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders may also be placed by calling
(202) 512-6000 or by using fax number
(301) 258-4066, or TDD (301) 413-0006.

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

For information on how to access GAO reports on the INTERNET, send an e-mail message with "info" in the body to:

info@www.gao.gov

United States
General Accounting Office
Washington, D.C. 20548-0001

Official Business
Penalty for Private Use \$300

Address Correction Requested

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100



The Function of Theory in International Relations

Author(s): Charles A. McClelland

Source: *The Journal of Conflict Resolution*, Vol. 4, No. 3, The Place of Theory in the Conduct and Study of International Relations, (Sep., 1960), pp. 303-336

Published by: Sage Publications, Inc.

Stable URL: <http://www.jstor.org/stable/172856>

Accessed: 01/07/2008 20:56

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=sage>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.

The function of theory in international relations¹

CHARLES A. McCLELLAND

San Francisco State College

During the past half-dozen years the interest in international relations theory has risen remarkably. The volume of writing, including descriptive and analytical materials, by the professional students of the field has grown to the extent that courses of study on theory may now be put on the sound foundation of a special literature. Collections of essays on international relations theory have appeared, and more will be published shortly. The development is clear and certain; therefore, it is timely to ask what this thrust toward theorizing intends and portends. Of what use is it? What place has it in the advancement of the science or art of international relations?

One may declare his belief or view on the matter and even find a place in one "camp" or another. Attractive statements on the function of theory for the field may be found readily. Thus Paul Nitze has stated the basic specifications of a general theory of international relations as

a relatively small number of abstract concepts which bear some continuing relationships one to another, an understanding of which relationships helps to illuminate and make more understandable the complex data comprised in

¹Revision of a paper prepared for the Symposium on the Place of Theory in the Conduct and Study of International Relations, The Center for Research on Conflict Resolution, Ann Arbor, Michigan, May 12-14, 1960.

the concrete world of international relations [14, pp. 1-2].

Kenneth Thompson, reporting the deliberations of a conference, wrote the following about the functions of theory in international relations:

First, and perhaps most basic, it makes possible the ordering of data. It is a useful tool for understanding. It provides a framework for systematic and imaginative hypothesizing. It gives order and meaning to a mass of phenomena which without it would remain disconnected and unintelligible. . . . The ordering of data can help the observer to distinguish uniformities and uniquenesses. . . . Theory holds out the tools whereby the observer can discover in the welter of events that which is recurrent and typical. . . .

Second, theory requires that the criteria of selection of problems for intensive analysis be made explicit. . . . Theory can serve to make more fully explicit the implicit assumptions underlying a research design and thus bring out dimensions and implications that might otherwise be overlooked.

Third, theory can be an instrument for understanding not only uniformities and regularities but contingencies and irrationalities as well [59, pp. 735--36].

In the same effort at identification, Edwin Fogelmann (32, p. 79) notes that the "uses of theory and conceptual frameworks may be described as follows: (a) they give coherence and significance to data and findings; they facilitate a true accumulation of knowledge; (b) they indicate areas for fur-

ther research; (c) they help alert the researcher to all relevant aspects of his work; (d) they may aid in prediction."

In general, it may be said that theory is intended to be a kind of yardstick for the evaluation of what already exists in the field of knowledge and a kind of signpost pointing into the future. Its preoccupations are, in other words, both with the Backward Look and the Forward Step.

This essay is an effort to go beyond the excellent statements noted above of the general functions of theory. There are serious problems and disputes which raise obstacles and obscure simple discussions of what theory is and what theory does (or may be able to do). It has been overlooked too frequently that the development of organized knowledge in a particular field is a social enterprise. In historical perspective it has been observed that, even in such austere fields as mathematics, highly individualistic and personalized contributions have had a shaping effect on subsequent development. John Rader Platt (45) recently has called to attention the wide range and diversity of talents which intricately complemented each other in the building of the present theoretical structures of the physical sciences.

The burgeoning field of international relations, with the great scope and variety of its phenomena, will probably be supported in the long run by even more intricately complemented efforts toward the development of competent theoretical frameworks. At this early stage it simply will not do for one man or one group to define and delimit international relations theory and to expect general acceptance of that formulation. Yet one of the important problems of the moment results from the large differences in conception of the meaning of the term "theory."

Some scholars declare that they have no

interest in theory and feel that the work they do is not hampered by the absence of the theoretical concern. Others seem to regard theory as almost any idea, view, or hunch about international affairs. There exists a conception of theory as the speculative and idealized counterpart of practice, making theory into a kind of systematic and persistent error. In the circles of political science and history, it is recognized commonly that theory may mean the history of ideas, the analysis of systems of political or social thought, the product of creative efforts to state or restate moral philosophies, or the formulation of strict schemes of explanation to account for observed events and states of affairs.

We appear to face in "theory" a series of problems akin in type to those encountered in the writings on Zen Buddhism. If we could say what theory is, we could probably give definition to the conditions and trends of a science or study of international relations, but we cannot. If we had one or several general theories of international relations, we would know what we meant by special and middle-range theorizing, but we do not. Research questions could be exploited systematically to fill gaps in our knowledge if there were guidance to choices according to a generally accepted theory, but this cannot be done at present. What can be done about such dilemmas?

It is one of the purposes of this paper to propose some pragmatic solutions. The first of these is the proposal that international relations be defined, for the time being at least, to include the efforts, current and past, to identify, characterize, and define the field and the body of relevant thought that has grown up around it. The impatient remarks of those who do not want to hear or read one further word on "theorizing about theory" become, under this definition, one of the elements in the body of thought

in the field, still open to description and analysis. As a point of view or attitude, such "anti-theorizing" may be traced through the four decades of constant examination, re-examination, and the examination of the re-examinations so characteristic of the study of international relations since it made its first appearance as an academic subject.

J. J. Thompson (50) once remarked that "theory should be a policy and not a creed." It is in this sense that the proposal is made to regard theory, in its broadest aspect, as the struggle to lay down intellectual policy for the field. Thus the definition of international relations theory is the total discussion to date of the conceptions of (a) what international relations is; (b) what the study of international relations is; (c) what is to be investigated, learned, explained, and taught; (d) what methods and techniques are to be used in the quest; and (e) all the ideas, proposals, and debates relating to these conceptions. It is the body of ideas of an extended conversation.

Theory-as-policy includes the literature of self-examination and evaluation, which, to date, has played such a large role in the history of the field. The catholic nature of the definition assures that other more circumscribed and "phenomena-centered" theoretical constructions may be included and accommodated. To keep the two meanings—that is, theory-as-policy and "phenomena-centered" theories—from becoming confused, "Theory" or "subject Theory" will be used in referring to the first variety and "theory" or "subject-matter theories" will be employed for the second.

The State of Subject Theory

With this much in hand, something can now be set forth in review of the place of Theory and theories in the current study of international relations. There are parts of subject Theory which appear to have

achieved general acceptance in the profession. Although they make only a short list, a number of Theory items seem to be accepted as settled concepts and, therefore, do not attract much discussion:

1. There are regularities and recurrences which are reflected in the concrete data of international relations. We make the assumption that these regularities and recurrences exist, that we can find them in the welter of the data, and that we can explain them by means and methods exposed to the public view. The epistemological requirement is that we make this assumption if we profess to understand. Faith in this presupposition is, indeed, the foundation of the social sciences.

2. Thus the general intent is to go about the business of finding out about regularities (and, therefore, by implication, about non-regularities as well). The study of international relations is to become increasingly systematic and analytic rather than only descriptive of certain categories of past events. This goal is shared no less by students of the persuasion of Hans J. Morgenthau who may believe that the laws of international politics are immutable and, therefore, that everything important to know has already been known than by some of the "radical behavioralists" who may feel that everything is yet to be discovered.

3. Although the study of international relations is not regarded as a universal subject,² it is the established view of the profession that any and all time-and-place spans may be used according to the re-

² The distinction I have in mind is the difference between a discipline which attempts to account for a single sector or aspect of the whole of human behavior or action (international relations, economics) and one which proposes to explain the whole of that behavior or action (sociology, history). But see also Wright (65, p. 5).

quirements of a particular piece of work. The roots of the field are deeply imbedded in historical studies, and history has been recognized repeatedly as an important resource. Our ultimate focus is on the phenomena of the here and now, however, so the data and meanings of the past are helpful and relevant to the extent that they bring the benefits of depth and proportion to the subjects in the prime focus.

4. International relations can be distinguished from other studies by the nature of the general subject it investigates. No other study seems to be concerned concurrently with the interactions between whole societies (or some analytic aspect of the whole such as the legal, the political, or the economic) and the relevant, separate actions issuing from each of these societies under circumstances of no superior human control system. Hence there is a perpetual problem of understanding and explaining the effect and the impact of the "domestic" on the "international" and the "international" on the "domestic" in an endlessly intertwined two-way traffic. There seems to be no perfectly simple way to state this "first idea" of the study of international relations, but it occurs, in one version or another, in almost all definitions and descriptions of the subject.

5. It is acknowledged generally that many disciplines and fields of knowledge contribute materials to the study of international relations, and, to that extent, it is interdisciplinary. If more than this is stated,

³The reasons for the overwhelming emphasis on contemporary affairs are not notably theoretical. Presumably, a scholar regarding himself in the field of international relations might spend a career analyzing and theorizing on the relations among the ancient Greek city-states without reference to or interest in contemporary situations. Such instances simply are now extremely rare.

the unsettled intellectual policy question of the extent and kind of interdisciplinary involvement is aroused.

6. A basic classification of the major subparts of the study of international relations has become established and stabilized. This is so to the extent that everyone understands broadly the references to international law, international politics, international economics, and international organization. There is no serious controversy concerning this classification scheme, although its status rests, quite clearly, on custom and usage rather than on logical grounds.

The area of agreement indicated above is quite small; modest as it is, one might expect that a representative committee from the field would accept such a list only after many changes and qualifications.

Even more difficult to name and describe are the intellectual issues upon which the scholars and teachers of international relations agree to disagree. It is scarcely possible to enter this arena except as one of the gladiators. The observations one may make about the importance or validity of a given issue are invariably to be challenged on grounds of personal bias or erroneous judgment. International relations has no clearly defined schools of thought; it has, instead, numerous individualistic conceptualizations about the "realities" and the study of those phenomena. Individuals sometimes find like-minded colleagues, but the mechanisms of professional social controls are so inoperative as to place every serious question at issue on a basis of personal acceptance or rejection. Whether or not this state of affairs is an indication of an undesirable condition of intellectual anarchy or a happy sign of the freedom of inquiry is, itself, one of the items of controversy. In arriving at a list of the more important matters in dispute in the Theory of international relations, one turns naturally to the literature of

self-evaluation and self-criticism of the field in order to secure a minimum criterion of choice. Such a list of significant intellectual policy questions might well include the following:

1. The relative weight and importance of subject versus subject matter.
2. The separate discipline versus the "synthesis of various special subjects."⁴
3. Normative approach versus empirical approach (and other aspects related to the "philosophy" versus "science" approaches).
4. Knowledge for its own sake versus the policy-relevant criterion.
5. Symmetrical theory versus common-sense theory.⁵

It is necessary to add, at once, a few qualifying comments. Some, if not all, of these matters can be approached tolerantly with a "let all flowers bloom" attitude. The black-and-white opposition indicated in the statements is an incorrect representation, since any student of the field can find certain intermediate, Golden Mean positions to advocate. There are, however, serious limits to the pluralistic outlook because standards of worth and workmanship are not readily sacrificed on the altar of generosity and tolerance. It comes about, then, that pluralism often operates as a mask, being defended in principle but subordinated in actual evaluations of the worth of the work of others (14, p. 41).

Second, it should be noted that it is the lack of "consubjectivity" ("consubjectivity" being a word used by Hocking and Brecht to refer to the common frame of reference, the common "reality-identification," or the

common understanding shared among many minds) that accounts in considerable part for the reappearances of these issues (9, p. 33). Hence it is highly probable that the five items listed above receive many different meanings and interpretations. Theory, attempting to explain itself into a state of coherence and consistency, probably is more vulnerable to confusions of meanings than to disagreements.

Lastly, the relationships among all five of the controversial questions named above are so close and so numerous that the treatment of one in isolation from the rest usually proves to be impossible. The discussion that follows of the Theory content in the "contested areas" is meant to illustrate the intertwined character of the issues and also to indicate why it is important to continue that extended conversation called Theory—no matter how futile it may seem—in search of the Utopia of unified or general theory. The writer is aware that some of the positions taken are partisan and contrary to the frequently expressed opinion that "the work of the field gets done by doing it and not by talking about it." This is, of course, a true sentiment, but so also is the view that the work to be done needs definition. One may take note of Alexander Meiklejohn's advice: "There is, I think, nothing in the world more futile than the attempt to find out how a task should be done when one has not yet decided what the task is." This is a suitable introduction to the first topic in dispute: subject versus subject matter.

It has been C. A. W. Manning who has called attention repeatedly to the importance of the distinction between subject and subject matter. In 1938, at the Eleventh Session of the International Studies Conference in Prague, Professor Manning, finding the deliberations of the conference becoming confused by the expressions of varying views on the normative question and on the

⁴ Fogelmann (32, p. 79) formulates this unresolved problem in the question: "Is international relations a distinctive or synthetic field?"

⁵ Again, Fogelmann (32, p. 79) asks: "What should be the relative emphasis on logical symmetry (or theoretical completeness) as compared with empirical relevance?"

possibility of one scholar-teacher addressing himself in a course to the whole subject, sought a simplification in these terms:

The question, "Is international relations a particular kind of subject," seems to me to be rather unsuitably posed. What we are given is not a *subject*, but a subject-matter; not the academic *subject* of international relations, but international relations in the sense of a particular aspect of the totality of social phenomena. . . . I begin with the subject-matter of human life on this planet, international relations being a facet or aspect of that life. I go on from that to a person called a student, who wants to understand that facet, and feels he understands it but imperfectly. . . .

So the professor is a student who is trying to help other students in their common task, which is to understand a little better the given subject-matter. I hope that that is not too vague, but I believe it is the true starting point. It is not necessary to enumerate in our definition economics, geography, history, political science, and all the other ancillary disciplines [66, pp. 228, 236].

By 1954 conditions had changed enough so that Manning's differentiation of subject and subject matter was introduced again but with the difference that he could report the emergence of a discipline, not of various offshoots of political science, history, psychology, geography, economics, and others, but of international relations per se. Manning quoted J. J. Chevallier's definition with evident approval and in recognition of the appearance of a new discipline.

Its concern is with a tangled intertwining of relationships arising, in all sorts of fields, among the various States within that special sort of "relational" milieu which is generally referred to as "international society" (as also, secondarily, between the States and certain so-called international organizations). This "international-relational complex" (*complexe relationnel international*) from out of which there are constantly cropping up those occurrences known as "current events," needs systematic describing and analyzing in the light of a fairly large number of kinds of knowledge, classed together

for this purpose, and to be antecedently acquired [38, pp. 10 and 56].

In the United States, at least, Manning's original attention to the subject-matter outlook now can be taken as a matter of course, and the issue of subject versus subject matter seems to have little point. Who would deny, for example, that certain of the aspects of geographical knowledge have important bearing on the study or understanding of international relations? The revival of the distinction has a different context, however, and it is related very directly to the successful emergence of the concept of the *complexe relationnel international*.

Two strategies for building a better *complexe* are open. One appeals particularly to those whose main concern is research. It most frequently consists of a single-minded concentration on the subject matter. The "outsiders"⁶ who have come into the area of international relations research from the so-called behavioral fields such as psychology and sociology are almost certain to confine themselves to theories of the subject matter or to "phenomena-centered theories" mentioned above. Why this tends to be so will be considered in a moment. The second strategy recommends itself to those who are concerned primarily with teaching problems. This strategy centers on questions of how the many segments of pertinent knowl-

⁶ "Outsider" may not be a fair term; geography, political science, history, and economics might all claim to be the original home of international relations, and the latter-day students of these fields have all studied the subject since its birth as an academic field. If there is a real difference, it is found in whether or not a person has studied the traditional or conventional literature and has adopted the working attitude and the style of discourse which are shared widely by students of the field. A person is "outside" when he does not show the "traits" or reveals his ignorance of the shared material.

edge are to be related, joined, organized, and interpreted. It is a subject-centered strategy, necessarily, and of importance to the younger generations of scholars upon whom fall most of the burden of undergraduate teaching and to whom the idea of a synthetic discipline or a levels-of-organization approach seems most appealing. Thus the exploration of how to translate and reshape some portion of the knowledge of social psychology (for example) to the particular purposes of the study of international relations may be regarded as an important concern.

The "research-centered" student, on the other hand, may have a difficult time in turning his mind to such problems. Even when he is teaching, the devoted research man with a predilection for theory tends to concentrate on such questions as: "How can I formulate a general proposition to cover this body of findings about events and occurrences?" and "Does my general statement say something that can be confirmed or overthrown through such and such research procedures?" These mental activities represent "real" theorizing to him, while the joining of slabs of knowledge from far-flung sources seems to be either inconsequential, impossible, or something other than theory.

The behavioral impact on the field amounts to a large step toward the realization of the ideal of the study of the *complexe*, but it also greatly extends the spread of an already expanded subject. How the subject can be held together in some kind of unity since psychological, sociological, and economic perspectives have been introduced is the contemporary version of the old Theory question, "What is international relations?" In contrast, the disciplines of psychology, sociology, and economics have come to a stage of development when their practitioners are no longer greatly con-

cerned by soul-searching questions about the identity, the legitimacy, and the total scope of the subjects. Instead, the major concerns in these fields run to "special" and "middle-range" theories for narrow spans of observed or observable phenomena and also to rigor in methods and techniques. The demands for operational definitions, for measurement, and for close delimitations of the areas of inquiry tend to crowd out the idea that a function of theory is to integrate and co-ordinate the whole body of existing knowledge. The behaviorist's preoccupation with subject matter is shown by Harold Guetzkow's conception of how the theory of international relations will grow:

The most useful theories will have to be, at first, small conceptual systems dealing with a restricted range of phenomena. . . . It would be fruitful to limit at first the predictions to minor international occurrences, rather than risking an attempt to forecast important global events. Confirmation and denials of parts of theories might gradually lead to a more firmly bulwarked system, eventually enabling the social scientist to predict more and more imposing events in international relations. . . . But is it possible to construct a theory of international relations with the extension of methods proposed? What form would such theorizing take? How would it be possible to construct small islands of theory, which eventually might be tied together into a more definitive theory-system? This article can not answer these questions but presents only some explorations [16, p. 426].

The conflict of subject Theory with subject-matter theories is removed when it is admitted that there is room for both to develop. The clash continues, however, when

⁷ Economics and psychology recognize the broader and the narrower scopes, however, in the labels "macro" and "micro" in economics, and "molar" and "molecular" in psychology. Except for the work of Talcott Parsons, "big" theory seems to be out of style in sociology.

it is implied, as in Guetzkow's observations, that the islands have yet to be constructed before they can be linked into a larger land mass. At the extreme, on the subject Theory side, is the assumption that the islands already exist, have been explored, and are at least partly joined together. A personal judgment that one is as important as the other as theoretical contributions—one being analytic and the other synthetic—does not, in the least, resolve the issue. Subject versus subject matter simply is a current item of disagreement in the field.

The unsettled question of the separate discipline or the synthesis of a number of special subjects has a long history of arid and inconclusive debate. Political scientists have begun, quite naturally, with the presupposition that the study of international relations is primarily a political subject concerned with official governmental relations among states. The desire for a wider focus has been expressed in the claim that international relations is really a special sociology. Georg Schwarzenherger's (49, p. 8) definition of international relations as the study of international society is well known. Alfred von Verdross, in a paper prepared for the Prague meeting of the International Studies Conference of 1938, declared:

The Anglo-Saxon expression "international relations" has multifarious meanings and is nothing else than international sociology. International sociology has to start from the results of the different national sociologies, but represents a science different from them. It applies sociological principles and rules to the society of States and other international communities [66, p. 23].

Norman J. Padelford and George A. Lincoln (43, p. 3) have written: "In the broadest sense, the field of international relations comprises myriads of contacts among individuals, business organizations, cultural

institutions, and political personalities of many different countries."

The separate and distinctive discipline existing on its own foundation but keeping its lines of contact and exchange open with other social science disciplines is a formula of reconciliation to some professional students but not to others.⁸ It is also proposed that fruitful borrowings and reinterpretations from many other social sciences will take place in the framework of the study of political behavior. Viewed in the perspective of the report of the Interuniversity Summer Seminar on Political Behavior (22, p. 1004), international relations might become free of its tie with political science because the study of political behavior is interdisciplinary and separate from the discipline of political science. Far removed from this outlook is Morgenthau's determined stand in behalf of a narrower kind of international relations study:

Today most institutions and students have turned to the study of international relations because of their interest in world politics. The primacy of politics over all other interests, in fact as well as in thought in so far as relations among nations and areas are concerned, needs only to be mentioned to be recognized. The recognition of this primacy of politics cannot but lead to the suggestion that among the legitimate predominant interests upon which international relations as an academic discipline might be focused international politics should take precedent [*sic*] over all others [40, p. 10].

Kenneth Thompson (58, pp. 440-43) has warned that, without a theory of politics to circumscribe and order the field of study, the result can only be an undesirable eclecticism.

The outcome of this argument over some

as Alfred Fernbach (13, p. 280) has noted: "It is from the special synthesis of the social sciences which international affairs embodies that its principles and methods of analysis emerge."

kind of separate discipline seems to be well on its way to settlement through practice rather than debate. Quincy Wright's judgment that a separate discipline is emerging through a synthesis of parts of a large number of other fields of knowledge has been accepted widely, if not universally (23; 32; 63, p. ix).

Far more controversial is the problem of what role is to be played by normative theory. Whatever the ultimate character of the discipline—separate, subordinate, synthetic, or unitary—the question remains of how "value free" it should be made. The conflicting answers generally have a bearing on the "pure knowledge" versus the policy-relevant issue. The two issues are often seen as one, although there is no logical imperative for such a linkage.

Normative theory for international relations involves the investigation and formulation of those sets of ideas which, if they were put into full practice, would be expected to lead to more ideal relationships among states, economies, societies, peoples, and/or countries. What goals should be pursued, how parts should be related and ordered to the whole of the international system, and how means and ends should be associated and realized are major concerns in normative theories. A normative theory is said to have achieved its maximum worth and clarity when it has been integrated fully in terms of a primitive value or values such as justice, freedom, God's Will, equality, etc. (9, pp. 119-24).

Empirical theory, on the other hand and in its "pure" form, is not concerned with the question of the "ought to be's" but only with the explanation of "what has been, what is, and what will be." It is a series of related propositions purporting to identify and explain, within a given frame of reference, either the data of observations already made or the expected findings from

new observations which the theory indicates could be made. The issue of international relations Theory, in this case, is the general policy question of the relative roles to be played by the two types of theorizing—normative and empirical. As timely as the latest book on the fundamentals of international relations, the problem has certain aspects which are commonly confused but need not be. It may be useful to specify which of the apparent difficulties can be eliminated from the controversy.

First, it is clear that the normative materials—social goals and values, "preferred outcomes," individual and group attitudes, norms, beliefs, and opinions—are as much open to study through empirical procedures as other subject matters. It is as appropriate and as reasonable to ask about what men value as it is to ask about what they do. A scholar may be disinterested in these subjects, or he may doubt, personally, that scientific investigations will succeed in identifying and explaining the value elements which are present in international behavior, but he ought not to insist that the approach is non-existent, wrong, or impossible.

Then, there are two arguments, sometimes raised for the purpose of destroying the distinction in question, which do not contribute to the solution of Theory problems: (a) that the formulations of normative theory are merely poorly phrased and mistakenly conceived statements of empirical theory and, therefore, can be eliminated (34), and (b) that behind all empirical theories and investigations are value assumptions which, in the end, deliver all theories to the normative fold.

Clyde Kluckhohn's observations on the interdependence of normative and empirical theories are particularly pertinent. He writes:

Existential propositions purport to describe nature and the necessary interconnections of

natural phenomena. Values say, in effect: "This appears to be naturally possible. It does not exist or does not fully exist, but we want to move toward it, or, it already exists but we want to preserve and maintain it. Moreover, we aver that this is a proper and justified want." . . . Finally, it should be noted that existential statements often reflect prior value judgments. In scientific discourse, at least, our propositions relate to matters we consider important.

"Nature" is one frame of reference; "action" is another frame of reference. In the former, one need only ask, "Is this the case (fact)?" In the latter one must ask both this question and, "Ought this to be the case (value) in the conceptions of the subject(s) of the enquiry?" The two frames of reference, as has been shown, are intimately related. . . .

Existence and value are intimately related, interdependent and yet-at least at the analytical level-conceptually distinct [44, pp. 892-94].

It is undeniable that normative propositions are offered frequently in loose or poorly defined terms just as it is so that we often become impatient with those prescriptive plans which lack a firm anchor in "reality" (speculative blueprints for "world peace," for example) (21, p. 185). Value judgments are present in the foundations of empirica theory, but also they are made throughout in all those operations which call for choices of inclusion and exclusion of hypotheses and propositions. We must admit that the hardest of "hard facts" does not always carry the same meaning and the same significance to all observers. To ignore the cognitive evaluating and the transactional processing of meaning-without fact from culture to culture, to say nothing of from individual to individual, would be one of the most serious defects from which we could suffer as students of international relations.

Yet, for all that, Arnold Brecht's massive demonstrations of the mistakes that have been made already in efforts to bridge what he calls the "logical Gulf between Is and

Ought" should stand as a warning to those who would seek a reduction and a final reconciliation of normative and empirica theory (9). As Brecht shows, normative theory can be aided and disciplined by empirica theory and research, but the line must be drawn sharply when we detect the tendency to establish the validity of one in terms of the other. The old debate in international relations over realism and idealism rested precisely on this dangerous ground. Stanley Hoffman's prescription appears to call again for the bridging of that "Gulf":

Our first problem is the clarification of values we would like to see promoted in the world—and as I have suggested, we cannot do so if we do not start with a view of man as, at least in part, a community-building animal, making moral decisions among alternative courses of action which all involve the presence of some values and the sacrifice of others.

Second, we must relate these values to the world as it is, far more closely than we usually do. A total separation of "empirica science" and "moral philosophy" would be disastrous. To go on repeating that only the latter can discuss what ought to be, that the former has neither the function nor the possibility of passing ethical judgments, since one cannot deduce an "ought" from an "is"—this cannot be our last word [21, p. 187].

Hoffmann reviews in brief the contributions of the empirica to the normative⁹ and arrives at a conclusion much in the spirit of the "realist-liberalism" solution that John Herz suggested several years ago (19). Hoffmann adds:

I would therefore like to suggest a task in which systematic empirica analysis and a philosophy of international relations would merge,

o It is not to be overlooked that an orderly way to treat various empirica theories is to consider them in terms of their different normative orientations. Claude adopts just this procedure to differentiate the concepts which underlie the approaches to topics such as regionalism, collective security, disarmament, and trusteeship (10, p. 12).

just as the empirical and the normative elements did coalesce in the great theories of politics and economics of the past. . . . We must try to build relevant Utopias [21, p. 189].

Hoffmann appears to overlook in his severe criticisms of past theoretical efforts—what he calls his "wrecking operation"—the continuous attempt to bring the normative and empirical elements of theory into a satisfactory relationship. The intellectual policy problem remains the same as it has been: to establish the effective connections between the *Is* and the *Ought* but without the erroneous and tempting operation which derives an *Ought* from the *Is*. For example, Hans Morgenthau's power theory, as it stands now after a subtle shift in direction, does, indeed, meet Hoffmann's requirements. When Morgenthau added the definition that national interests are to be defined in terms of power (41, pp. 3-13), his meaning was commonly taken to be that the end of power-seeking and power-holding is power itself. But Morgenthau, by modifying the idea of the "permanent" and "objective" character of national interests (41, p. 9), has opened the field to the admission of normative elements of "a relevant Utopia." Power may be regarded as an intermediate value which, when properly "maximized" in its setting, may advance more basic or primitive values as specified by the Utopia. We can dislike the theory or we can demand that the identification of power be improved so that its meaning is enhanced (11), but we need not imagine that anyone has disposed of the theory once and for all. If we regard "national interests" as so many blank spaces to be filled in from a "realist" philosophy (or, better, any one of several philosophies) of international relations, then the power theory will fit Hoffmann's prescription of togetherness for normative and empirical theory. Similarly, Hoffmann's strictures heaped

upon the "policy science" approach of Harold Lasswell are undeserved (21, pp. 10-12 and 188). Lasswell has been doing all along (12) what Hoffmann says we must begin. In 1956 Lasswell observed:

Our first professional contribution, it appears, is to project a comprehensive image of the future for the purpose of indicating how our overriding goal values [these are set forth many times and in many places by Lasswell as human freedom and the dignity of the individual] are likely to be affected if current policies continue.

A closely related contribution consists in clarifying the fundamental goal values of the body politic. . . .

The third task is historical and scientific. It is historical in the sense that by mobilizing knowledge about the past we are enabled to recognize the appearance of new patterns and the diffusion or restriction of the old. It is scientific in the sense that we summarize the past in order to confirm (or disconfirm) propositions about the interplay of predisposition and environment [31, p. 978].

These observations on Hoffmann's recent critique have been made to illustrate what is *not* really at issue in the topic of normative versus empirical theory. In the first place, since the beginning years of international relations as an academic field, one viewpoint has persisted on the theme that the subject must combine normative and empirical elements. The nature of the combination has been, of course, a subject of continuing debate. The current issue, however, is concentrated in the question of whether or not the normative element should be reduced to a minimum level, comparable to that found today, for example, in psychological theory and research. The choice should be approached not in terms of personal taste but with reference to what functions the study of international relations should perform in the future.

If our discipline and, along with it, our theory are to be preoccupied with the preparation and presentation of advice to pol-

icy-makers and publics-if it is to be mainly of the breed of the engineering sciences-then the development of the best possible normative theories is a main task. On the other hand, if the prime objective is to ascertain how international relationships are changed and maintained, to find out what takes place in international behavior and why it happens one way and not another, then empirica} theory is the matter that requires our most serious effort. One objective does not preclude absolutely the pursuit of the other, but the problem certainly does involve questions of emphasis and proportion. If the search for knowledge for the sake of knowing more through empirical observation and interpretation became the prime task of the discipline, normative studies would still be helpful as an aid in retrieving the thought of the past and in disciplining those evaluative assumptions and judgments which are, indeed, inescapable. A place would be secured for the normative content at the important but modest level that Kenneth Waltz (14, p. 67) has indicated: "The function of political philosophy is to help to form, sharpen, and critically ground the fundamental understandings that we all build up somehow in our minds."

For the sake of completeness, two further observations should be added, although each of them calls attention to an obvious condition. It is possible to pursue knowledge (or, perhaps, it is wisdom) for its own sake in the normative frame of reference. The link of the normative with the policy-oriented concern is not essential. This is true in one direction but not in the other: a policy inquiry into "what should be done" must give consideration to both the Is and the Ought. One of the chief reasons for the criticism of the policy science approach may be the suspicion or belief that the study of the Is is carried out inadequately-it is slighted and distorted in deference to

the Ought concern. An abuse that is too often seen, under license of the "policy science" approach, is that of the academic man who plays a Monday-morning quarterback role and who, from a self-assigned position of superior wisdom, constantly writes and speaks on what the Administration should do or what it should have done in foreign affairs. The reference is not to carefully worked-out means-ends analyses of foreign policy; the object of criticism is the familiar figure of the academic commentator who airs his personal or political preferences on virtually any passing item of world affairs with a tone of authority suggesting that his conclusions are derived from intensive scholarly research.

The fifth item on the list of the issues in dispute in international relations Theory-symmetrical theory versus common-sense theory-is a very recent addition to the discussion. Furthermore, it is concerned only with empirica} theory. Common-sense theory is usually a descriptive and generalized explanation of something anybody should be able to discern in the phenomena of international relations. Hobson's explanation of imperialism is an example (20). To account for the late-nineteenth-century expansionism called the "new imperialism," Hobson described certain conditions prevailing in the development of the capitalistic economy, the special groups that stood to profit from high returns on investments in new colonial territories, the machinations of these groups to secure popular approval, and the resulting involvement of the government in support of these enterprises. The theory has high plausibility. Things *could* be conceived to happen as Hobson explained. The remaining question was: Did these occurrences and developments take place as explained? William Langer was able to show from history that the theory was incorrect: the placing of foreign invest-

ments and international political action did not work in tandem closely enough to support Hohson's theory (30). The theory and the disconfirmation were easy to understand because the references of the ideas and arguments were familiar-"commonsensical."

Symmetrical theory is more difficult to comprehend because it departs so readily from straight-line explanation. It will seek to account for all possible courses, alternatives, and outcomes in a hypothesized state of affairs. Further, the state of affairs may be entirely "theoretical" so that the theorist may take pains to indicate that his structure of related propositions does not have a reference to anything that has ever happened in history or, perhaps, ever will. In effect, the theorist makes his way into all the nooks and crannies of a conceptual structure built in his "imagination" until he gets all essential parts in the right places and in the right relationships. The result may be a symmetrical theory. An example in the current literature is Morton Kaplan's *System and Process in International Politics* (24). No small amount of the puzzlement and irritation that greeted this book can be attributed to Kaplan's concern to construct symmetrically. Just why the "something" named "The Universal International System" has just five rules, no more or less, governing its behavior, particularly in view of Kaplan's declaration that it has never had a historical counterpart, would seem to be mysteriously arbitrary (5; 24, p. 47). The theory of which "The Universal International System" is a part cannot be judged by the "facts on file," since there are no facts to which a large part of the theory might correspond. There is no way to determine the worth of such a symmetrical theory against the yardstick of common sense or surface plausibility. Only by going through the same mental processes and the same

mental experimenting employed by Kaplan might we be able to decide if he has done his work completely and properly down to the detail of five rules, no more and no less, for one of his theoretical systems.

The history of science shows many instances of the usefulness of theories which violate common sense. This has been so even when the theory contains propositions about phenomena that could not possibly exist in nature. The idea of a "perfect gas" is an illustration. The answer to this, quite naturally, is that international relations is not physics and that what is useful in one field is not necessarily applicable in another.¹⁰ There remains, nevertheless, a real issue of the degree to which theory in international relations should have a *direct* fit and correspondence with the data. The choice between the types illustrated by Hohson and Kaplan cannot be made easily. Kaplan's reasons in behalf of the symmetrical type are not to be dismissed out of hand:

The analysis of systems without historical counterparts has definite value. In the first place, the models of international systems with historical counterparts contain predictions that new kinds of international systems will arise if certain conditions hold. Therefore a statement of the characteristics of international systems without counterparts is necessary if models of existing systems are properly to be subject to confirmation. In the second place it is desirable to make predictions about how such interna-

¹⁰ Kenneth Bock (6) has written at length to warn against: the dangers of transfer of concepts by analogy from field to field. Arnold Brecht (9, p. 523) notes, however: "Analogies are frequently instrumental for the birth of hypotheses. . . . The analogies, e.g., that exist between the game of chess and the game of politics are not in themselves 'theories'; only if used to 'explain' political phenomena do they become 'theories,' and then they serve as no more than one item, a 'hypothesis' within the 'theory.' "

tional systems will behave if they do arise. Unless this is done, predictions concerning the transformation of existing systems will be too loose for proper confirmation. . . .

If theorizing stops-rather than starts-with overly simple models there will, of course, be no progress and even no really operational knowledge. The real test is: Do the simplifications aid progress in research, or do they obscure important relationships and thereby detour science into the study of interesting puzzles? [24, pp. 21-22].

How the "real test" can be made in time Kaplan does not say. We have here a case of an unsettled Theory question: against what criteria shall we judge and measure theories?

The Case for Improved Theory: Criticism and Classification

Some consideration of the present content of international relations Theory, in the categories of both settled and disputed intellectual policy questions, now has been given. A full exposition of the rest of the specific content, including an inventory of relevant normative and empirical theories, would require much more space. In any case, there is no place to which one may go to find out exactly what theoretical materials exist to be included in an inventory. The immense task of digging out, sorting, and arranging the ideas of international relations has never been attempted seriously.

It would be a grave misstatement to declare, however, that the field or discipline of international relations does not have, at the present time, any command over its conceptual resources. Any seasoned student of the subject can draw up immediately a list of explanatory concepts and meanings that have been advanced and used to interpret the specific and concrete events and developments of international relations. What comes mostly to mind are separate thought items most easily remembered by

men's names or through association with ideologies, with perennial topics, with historical periods, or with recent controversies. In a stream of recall from stored information we might find: Plato's explanation of war and militarism as products of the rise of specialization and the division of labor in the economy and society of a country, Machiavelli's advice on how to deal with conquered peoples, the Manchester School's Utopia of a warless world of interdependence and mutual understanding built from the practices of free trade, Lenin's doctrine that war and imperialism are the necessary phenomena belonging to the last stage of capitalism, Kennan's tactic of patience and delay in foreign affairs until the incipient internal collapse of the Soviet Russian regime should take place, the frustration-aggression hypothesis of the sources of social and international conflict, Gandhi's non-violence teachings being carried now into discussions of disarmament techniques, the traditional Chinese and Japanese images of international relations cast in the mold of hierarchy as experienced in the authority and loyalty patterns of the family, Rostow's notion of the "take-off" stage of economic development, Amster's original formulation of the mechanisms of mutual deterrence under thermonuclear conditions, and the teachings of Christos Asoka, or Schweitzer-extended through the ideals of love, altruism, and self-subordination to the reform of man's relations to man.

The listing could go on and on to illustrate that the accumulation of the interpretations of the Is and the Ought relevant to understanding and explaining international relations is very large. It is also evident that these vast resources of ready-to-use meanings have an immense variety. The differences in Kennan's and Gandhi's "policy prescriptions"-of what ought to be done in real-life situations-are obvious and do not

require comment. In the great crowd of ideas that are at our disposal when we set out to assign meanings to some series of tangled events—for illustration, we might think of the recent relationships of the French and the Algerians—are explanatory concepts as broad as the idea of national self-determination and as narrow and time-bound as the sea-wall notion of empire: if a nationalist uprising is allowed to succeed in one part of the empire, that will be the signal for new troubles and renewed nationalist disturbances elsewhere. Mixed together in the pool of existing concepts hearing on international relations are pessimistic and optimistic doctrines; individual-centered and group-centered conceptualizations; "truth assertions" of the most varied kinds about the nature, the sources, and the workings of motives; numerous notions about forces, influences, and factors in the relations among states and peoples; and such partially obscured philosophical assumptions as those noted by Arnold Wolfers in his remarks on the Anglo-American faith in the human ability to control international relations contrasted with the Continental belief in the necessity to submit to and endure the consequences of forces in international relations beyond men's control.¹¹

Is there anything that can be done with this extensive agglomeration of idea, thought, doctrine, concept, and interpreta-

tion to make it serve better the systematic study of international relations? Should some method of control exist so that we could quickly determine whether or not the "newest and latest" explanations of international phenomena are merely restatements of old and well-explored concepts? Or would it be wiser to cast out the entire lot according to the judgment that such "folklore" is subscientific and, therefore, not to be trusted in a framework of reliable knowledge? Would it not be better to start anew with controlled observations of international behavior and to derive from these results only the generalizations that are genuinely confirmable?

The answers appear to have already been given to this line of questioning. The professional students of international relations are not inclined in the least to turn their backs on the past and to renounce the knowledge and wisdom found in history. No possibility exists to begin again with a clean slate. Instead, there is an emerging synthetic discipline built on intellectual contributions, past and current, from a number of other fields of knowledge. This discipline cannot be expected to become competent if its practitioners cannot solve the problem of how they will select and order the ideas and meanings they have borrowed, however. International relations must be only a special field dealing with only a sector of the total of social reality. Hence it must have built into it the ability to discriminate between what it can use and what it cannot. Since the study of international relations must consist of a selective mixture of old and new materials—both data and concepts—there must also be standards of judgment accepted and respected by the students of the field. But these standards, as was noted above, are precisely what we lack. There is no unified theory; there is no general acceptance of any broad approach or of any discriminating generalized expla-

¹¹ The alternate forms are given by Wolfers (61, pp. xix-xxii) as the "philosophy of choice" and the "philosophy of necessity." An illustrative statement of the philosophy of choice and its particular bearing on a basic subject of international relations is made by Claude (10, p. 14): "International organization has been built, consciously or not, upon the assumption that nations are not prisoners of destiny but reasonably free agents. Its theory is that the relationships of states can be modified, even while the present system remains unchanged in its fundamentals."

nation which would tell us how to pick and choose.

It is a Theory question to ask how we may escape from such an unsatisfactory state of affairs. It is, in fact, a first-priority task of the Theory of international relations to attack a dilemma which would require at the moment either an authoritative statement, in the name of the profession, of the scope and competency of the study of international relations or a prolonged wait in limbo for the appearance of a convincing general theory. What can he done that is not already being dane adequately and sufficiently?

Two mutually reinforcing kinds of activities would nurture and promote the emergence of an emerging discipline. Both activities helong in the category here characterized as subject Theory. The flrst recommendation is that we, as a whole profession engaged in writing, editing, reviewing, advising, training, and teaching, greatly increase the quality of our critical evaluations of the *thought* in the study of international relations. The second proposal is that a system of classification for the propositions, truth assertions, interpretations, and theories which purport to explain the phenomena of international relations he developed.

The need for intelligent criticism is common to every organized academic field. The study of intemational relations presently is in a peculiar position of having the advantage of sound criticism in all aspects except the ones most vital to its development. Without special bidding, the economists, historians, sociologists, and other specialists evaluate the works on international relations in terms of the standards, interests, and needs of their own disciplines. But, on the other side, the specialist in international relations usually neglects to direct his attention .to how a new study does or does

not contrihute *to his own discipline*. The historian will concentrate on the value of the historical content of a research report or a new study which lies somewhat outside the normal range of historical writing. He will consider whether or not something has been added to the historian's understanding and knowledge, and, usually, if his evaluation is in the form of a review, he will make a specific judgment on the worth of the study *as history*. Ordinarily, the international relations student will approach a hook on diplomatic history (let us say) in its own frame of reference *cis history* and will ignore largely the question of its specific contrihution, if any, to a discipline of international relations. The practice of neglecting the conceptual aspect-the theoretical component of international relations-is easy to document. Because the point may be misunderstood, if for no other reason than that it has been discussed hut little, a specific illustration will he given.

Henry Kissinger is a well-known younger scholar in the field of international relations. His mast prominent work to date is *Nuclear Weapons and Foreign Policy* (27) -presumahly, a title as central to the discipline of international relations as anyone can imagine. The place to which we naturally turn for an evaluation of such a work, in the terms of *our own field*, is *World Politics* hecause of the extensive review essays that appear in that journal. William W. Kaufmann contrihuted the *World Politics* review of Kissinger's hook (25) . We anticipate, as a matter of course, a full criticism of Kissinger's hook as either a major or a minor contrihution to our field of knowledge.

Now Kaufmann's review is nothing if it is not severe and searching. If we were students, primarily of military affairs or even military staff officers, we would find Kaufmann's ohservations of great interest and

relevance. It is a fine review for civilian observers as well. But that is just the point. Kaufmann has just one sentence in a long essay showing that he is aware of the field and its students: "This terrible revolution [the advent of nuclear weapons] and the constant effort somehow to come to terms with it have left their mark on the students of international relations as well" (25, p. 579). That is all. We do not even find out what the "mark" is or what it portends. Kaufmann, one may be sure, would argue that the orientation of his entire discussion provides an implicit and obvious connection with the interests shared among most if not all students of contemporary international relations. But can the matter be left at the implicit level? Is the connection so obvious? This writer thinks not; such subjects as military policy and military technology belong to a well-rooted discipline of military science and have their own functions within that field, both practical and intellectual. But the bearings, the relationships, and the contributions of these subjects in the study of international relations are far from being identified, placed, and determined. Why does a member of the discipline as meticulous and thoroughgoing as Kaufmann fail to treat the theoretical underpinnings of Kissinger's work? Clearly, Kaufmann is one of the subject-matter-centered specialists. It is a fair guess that it did not occur to him to consider the basic assumptions and theories underlying Kissinger's analysis, although this might be done readily in connection with Kissinger's previous work (26). Just where and how the studies of military strategy and policy connect with a body of principles of international behavior would be a central concern of a critic interested first of all in the discipline of international relations.

We, in international relations, have not developed sufficiently our critical skills in

the handling and evaluation of the conceptual materials which are used in attempts to explain phenomena in the particular, peculiar frames of reference at the root of the discipline. The habit of looking at any new work and of asking, as do the historians and economists for their respective fields, "What exactly does this contribute to the advancement of knowledge of the field?" needs reinforcing. There are exceptions, of course. In the same issue of *World Politics*, a few pages beyond Kaufmann's review of Kissinger, there is another essay by Ernst Haas (17) that asks the question and seeks the answers with respect to the theoretical and systematic contributions of several books to the knowledge of an international relations subject, namely, international organization. Haas is one of the few scholars in the field today who has systematically sharpened his critical skills as a theorist of international relations.

In the end, it may be necessary to recognize and develop a new specialization. Perhaps, we need even now a number of "full-time" theorists who will provide a steady supply of theoretical analyses and criticisms of the work contributing to the discipline.

A chief cause for the relative lack of competent theoretical criticism was noted above. In a sense, we do not know enough about the general "mission" of the study of international relations to determine what contributes to it. Given all the existing circumstances, the progress that can be made most immediately may come best from the building of inventories and classifications. This may appear to be humble work. It may seem that a call for a comprehensive classification scheme conceals a secret yearning for a new *Summa theologiae*—and a demand for intellectual closure—or, on the other hand, it may be taken as a further indication of the way the searchers for em-

pirical! theory sublimate their drives when they begin to find out that they cannot do what their programmatic declarations have specified (23, p. 94). There is a simpler and more practical viewpoint. By efforts toward the building and filling out of inventories and taxonomies, some steps can be taken toward correcting the paralyzing relativism which makes every debated issue of Theory merely a matter of personal preference or private belief. International relations takes in a span of subject matter too broad to be strait-jacketed within one simple set of organizing concepts or beliefs which somebody decrees shall be accepted. "Multiple realities multiply perceived" is the belief that most appropriately applies to a classification of ideas and meanings of this field.

There is, then, a task of taking inventory and classifying the meanings that are employed. Surprisingly little energy has been invested so far in these enterprises, but what work there is has been recognized and used. Frank Russell's pioneering effort, *Theories of International Relations* (46), is to be noted. The two-volume *Study of War* (64) stands as a monument of inventory and classification not only for its central subject but also for international politics in general. The advent of the nuclear-missile age calls mainly for additions to that great work. Quincy Wright's *A Study of International Relations* (63) has turned out to be most useful as a textbook and seminar guide because it is an encyclopedic review of the ingredients of international relations and their combinations in various conceptual frameworks. There are, of course, some valuable inventories for special subject, such as Klaus Knorr's *British Colonial Theories* (28). The collection of readings done by Arnold Wolfers and Laurence Martin stands out, on the other hand, as a contribution,

in part, because there are so few other available resources of related character.¹²

And, on the horizon of work in progress, there is much too little to report. Myres McDougal proposes an analysis of the concepts of "the diverse systems of public order" in the contemporary world (37). Other inventory projects for international law are afoot also. The International Relations Program of Northwestern University and Denis Sullivan are preparing a "propositional inventory" of the general meaning statements found in the introductory textbooks on international politics. There is nothing in being or even planned such as the study of the fundamental ideas and conceptual structures of basic sociology that Hornell Hart has been doing for that discipline. We need examinations and studies which would make explicit and available the intellectual content that is applied to the data of reported events.

Elsewhere, this writer has attempted to review some of the problems of classification (36) in international relations and to make a few suggestions regarding the development of classification systems (35). The rationalization of these interests stems from the strategic idea that order and coherence in the study of international relations can come more from attention to the "control point" where interpretations and meanings are gathered, criticized, analyzed, and constructed into systems than from inspecting the stream of concrete data. Initial attention should be given to the present means of classification and discrimination.

¹² The authors note: "If specialists in international politics with rare exceptions have neglected political theory, the political theorists in turn, departing from older tradition, have paid little attention to what the thinkers of the past - Machiavelli not always excepted - have had to say on international relations" (61, p. ix).

How do we categorize and classify at the present time?

There are two levels of categories in common use. The first was mentioned above: international politics, international economics, international law, and international organization, along with outriders of more dubious or uncertain status—foreign-policy analysis, international communications, military policy, decision-making analysis, institutional studies, world law, comparative foreign policy, intercultural relations, international economic development, etc. This is a first order of discrimination for the "object-slabs" of the field. When one pushes on to a second level of greater particularity, there is found a large and somewhat indeterminate collection of research interests and teaching topics. No attempt will be made here to catalogue this content; a few items among many may be mentioned to indicate the type of classification that is used: the state system, nationalism, imperialism, the peaceful settlement of disputes, world federation, economic integration, balance of payments, geographical factors in world affairs, diplomacy, collective security, national character, ideological factors, major problems of international politics, etc.

At this second level of classification it is not so much the case that each item is isolated in its separate category as it is that the conceptual connection between items is loose and uncertain. Fred Sondermann's note on the "box-car" approach to the coupling of topics in some textbooks reflects the point nicely (54, p. 102). Our main stock in trade happens to be such categorized items which we usually recognize as more or less "standard topics." In addition, we commonly utilize another classification device to lower further the level of generality of concepts and encompassed

data. It is the simple gridwork of time-place distinctions (36, p. 226).

This is about the extent of the classification structure for the over-all content of international relations. The first improvement toward greater precision, closer integration, and increased utility for this structure of classification would be an addition at the first level of a major category called international relations theory. Subcategories may be named, in the perspective of this paper, according to the main functions already noted: (a) criticism of concept formation and concept applications; (b) collection, co-ordination, and classification of the whole body of interpretative and meaningful ideas used in the subject, past and present; (c) creation of normative structures; and (d) discovery and formulation of empirical conceptual schemes, models, and partial theories. These four divisions of international relations theory include what was called subject-Theory and subject-matter theories for the Backward Look and the Forward Step.

To propose an addition to the structure of classification is one thing; to consider a modification of the structure itself is quite another. It must be admitted at the outset that any plan for overhauling the mental processes and habits of categorization in the minds of the participants in an intellectual enterprise as large as the study of international relations is unrealistic. A cardinal principle may as well be set forth: if we are going anywhere new, we shall start, certainly, from where we are. The two levels of categories are going to remain very much as they are now in actual practice.

The addition of a major item—international relations theory—on the first level has been proposed. There remains to be suggested one further addition of a classification scheme *between* the two existing lev-

els. That is to say, we need to build a certain type of discrimination lying between general subjects and specific topics and in such a manner that the intermediate classification extends to both of the existing sets of categories. Why such an addition is needed and also why international relations studies face a special problem in this regard require an explanation.

Most social scientists have now become familiar enough with semantic principles to be conscious of the differences of function of general and specific statements. That the object-referents of signs and symbols ought to be kept in mind is a kind of ordinary working formula. Thus the level of generality of a given sentence is a matter that we work out with little difficulty. Which is the more general of two statements: (a) "As soon as the Russian mobilization was ordered in 1914, World War I became virtually inevitable" and (b) "The maintenance of peace in international relations depends on the abilities of the major states to wage a war" is an easy question to answer. The Russian mobilization statement is specific because we can fix the when, the where, and the who of the situation. The *para bellum* assertion, on the other hand, is highly generalized.

Is one of the two types "better" than the other? From the standpoint of methodology, there are two answers. The first is the observation that which is "better" depends on the use in the same way that a scalpel is better for surgery than a bulldozer, but the latter is better than any scalpel for moving mountains. The second answer notes that in the social sciences one type of statement can be made complementary to the other so that their relationship becomes a mutually reinforcing one. The goal is to get the two types of propositions to work together so that additional "true" specific statements (ideally, products of minimum inference

drawn from direct observations) build increasing confidence that the general, covering statement is "true."

Now, what has just been stated is nothing more than a copybook rule about the method of systematic generalization. It refers to a normal procedure in empirical theory-building. The important consideration, however, is the difficulty of choice between two routes of generalization from the particular to the general. Other social sciences may face this problem, but it is, apparently, more crucial in international relations than in other fields. *The choice is between a "culturally rooted" mode and a "supracultural" mode of generalizing statements about international behavior.* The conceptual orientation of the generalizer may be such that his structure of explanations is intended to apply only to the actions of a certain group of culture-sharing people. Or it may be such as to imply that the generalized meanings are meant to apply to all men because of a facet of human nature or because a certain item or trait is common to all cultures or has spread and become common in all cultures. These choices are not merely preferences for higher or lower levels of maximum generality. The distinction being drawn for the purpose of indicating the need for a further classification of concepts can be shown best from the two statements about war.

If I should assert: "As soon as the Russian mobilization was ordered in 1914, World War I became virtually inevitable," the significance of my statement depends on references to other meanings I have not included. Do these other references—whether more general or more specific—pertain to characteristics, or situations, or previous experiences of the Russians, and, in this instance, of the neighboring Austrians and Germans? Is my focus, in other words, being held to the level of meanings

about how Russians act and Austrians act and Germans act, either separately or all together? Or is it the case that, in making the assertion about Russian mobilization, my frame of reference was adjusted to regularities in international behavior which bold, presumably, without regard to cultures and nationalities? Do I intend, in other words, to say that the Russian mobilization in 1914 illustrates that, whenever the war plans of states belonging to opposing coalitions reach a certain stage of development, the information that one war plan is being put into action will automatically activate others? The two structures of generalization are quite different and involve different modes of organizing knowledge.

In order to remove any notion that it is only the degree of generality that is involved, let us consider the possible alternate references of the other statement: "The maintenance of peace in international relations depends on the abilities of the major states to wage a war." The frame of reference may differ, as in the preceding case. I may mean that it is a human characteristic in any individual and any group to fear being hurt, so that mutual military threats cause all parties to retreat from situations which are known to lead to organized physical violence. The possibility is, however, that I have in mind the particular historical-cultural state of affairs of the period since 1954 in which the Russians and the Americans not only share the means for waging highly effective war with nuclear weapons but also have developed, in each country, a "cultural trait" of believing that the presence of the new military power "deters" to the point of preventing any outbreak of war between the two states. Let it be noted from these illustrations that being very specific with regard to historical fact does not, alone, clear up the problem.¹³

There are, as the preceding exercise

seeks to demonstrate, alternate modes of generalization—one that is rooted in particular cultures and another that is above cultures or supracultural. Confusion arises in theory from the failure to make explicit the mode that is employed. It would be sensible to clear up some of the ambiguity and the misunderstanding in our professional discourse by becoming aware of the problem of references to other, related propositions. The matter is important enough so that the systematic arranging of theories and propositions according to their types—either culturally rooted or supracultural—is warranted. The situation reduces to this: we are sensitized to *semantic* problems of the links between symbols and objects, but, in a field where there are at least two major possibilities in the mode of explanation, we have been too little aware of the *syntactical* problems which concern the links of symbols to other symbols. This shortcoming indicates the need for correction, and the habitual use of separate categories for the two modes would do something to rectify the situation.

One additional circumstance in the study of international relations creates a special problem. From time to time the problem has been brought to the surface. In the Prague meetings of 1938, for example, Ludwik Ehrlich declared:

Is the study of international relations the study of world affairs?—and my answer to this first part of the question is in the affirmative—or is it the study of relations between States? On that point also my answer is: Yes. I he-

¹³ Some may find that the distinction being made is too delicate. The writer may be guilty of a foible; on the other hand, how often do most of us try to "prove" generalizations by citing specific events of history (low level of generality) but without considering whether these events purport to demonstrate "cultural" or "universal" regularities?

lieve, in fact, that the question presents itself under those two aspects [66, p. 238].

In 1954 Richard Snyder wrote:

There seem to be only two ways of scientifically studying international politics: 1) the description and measurement of interaction; and 2) decision-making- the formulation and execution of policy. Interaction patterns can be studied by themselves without reference to decision-making except that the "why" of the patterns cannot be answered [53, p. 43].

In 1957 Fred Sondermann remarked:

A review of these four textbooks emphasizes one pressing need: students of international relations should become more precise about the nature of the difference between foreign policy and international relations. . . .

One must start with the assumption that "foreign policy" and "international relations" are not identical! (hence two different terms); although they are more often than not treated as if they were synonymous and many texts in international relations turn out to be texts on the making of foreign policy [54, p. 108].

These comments, in different ways, get at the problem of the perspective of the observer. It is difficult to say how much bitter controversy has arisen from the lack of a common understanding that there are two or three fundamental ways of looking at the phenomena of international relations, but there has been much, certainly. These different ways are closely related because all of them focus on the same general range of subject matter, and they constitute only a division of labor built from different emphases. For these reasons there should be a minimum of confusion on which is which. Conscious use of classification serves this purpose.

One perspective may be called the "actor focus"; Harold Sprout refers to it as "policy analysis."¹⁴ The actor focus causes international relations to be considered, as a whole, as the sum and product of all the policies pursued by the separate actors-be-

they states, economies, nations, or other entities that may be designated. Attention is drawn to those who make decisions in the name of the collectivity called the "actor," what the decision-makers decide to do, how the decision-makers for other actors respond, and what is decided next.

Theory and research in the "actor focus" tend to concentrate, very properly, on the processes that occur within each "actor" and that feed into the making of decisions. The "natural logic" that ends-means analyses are the best methods for studying international relations appeals to the student habituated to the use of the actor focus. Studies carried on for the purpose of directing or correcting a country's foreign policy fall easily into this perspective, and "policy science" carries a heavy emphasis in the same direction.

The "interaction focus" draws on quite different interests and procedures. Its devotees are unlikely to be interested centrally in policy problems mainly because they take an "outside" viewing position. Equilibrium analysis makes more sense than ends-means analysis to the student of interaction patterns, and the normative problem connected with identifying goals and purposes may become subordinated. Or, if this is not the case, the normatively inclined theorist has a wide canvas upon which to paint in the lines of the ideal patterns of international relationships. Pattern and con-

¹⁴ Sprout's assessment of the most basic perspectives for the study of international relations is not unlike the descriptions in this essay but with a difference that Sprout holds closely to his concern with statecraft, while I see statecraft as one among several other phenomena of equal or comparable importance to the study of international relations. Sprout writes (55, p. 9): "Basic research clusters mainly around three foci: (1) foreign policy analysis, (2) capability analysis, and (3) international system analysis."

figuration in the relationships, themselves, are main objects of attention. The perspective of interaction comes "naturally" to students who are attracted by problems of form and *How* and of related questions of fluctuation and change.

Snyder suggests that the interaction approach cannot answer the "why" questions. This writer suggests that he may be mistaken in part; an interaction inquiry will be unable to answer the same questions as decision-making, but the patterns of international interaction are open to causal explanations through the consideration of patterned conditions and characteristics inhering in the various actors. It is this lead that is followed theoretically in considering the diHering impact of highly "modernized" nation-actors and "emerging" nation-actors on the patterns of international interactions (33; 36, pp. 241-46). The studies of social modernization are progressing so rapidly that the opportunity to consider empirically the question of this impact on the *Hows* of international interaction increases almost daily.

The difficulty created by the failure to discriminate between the actor focus and the interaction focus is demonstrated most vividly when two men undertake a collaborative task and are unaware that one conceives the field in one perspective and the other in the other. Until the discovery of the differences in basic outlook occurs, nothing seems to go right. What one sees the other cannot; what is highly significant to one appears inconsequential to the other; and the procedures and analytic operations of one will seem dead wrong to the other. In the resolution of the problem the follower of the actor-focus approach will be likely to hold stubbornly to one last objection. He cannot see how the interaction viewpoint can be maintained in light of the fact that there is no place for international

interaction to occur except *within* the experience of the actor. What is the location of the interaction? This is a false difficulty. It may be disposed of by the consideration of a direct analogy. The location is in abstract space in the same way that the "business cycle," with its patterns and characteristics, has a "place" to occupy despite the fact that the actual experience from which it is derived is distributed among numerous actors-investors, producers, consumers, etc.

The actor focus and the interaction focus on international relations are distinct perspectives, and the one may turn out to be as useful as the other to the field. A third focus has been identified, particularly well, by Harold Sprout (55, pp. 9-10 and 15). He calls it "capability analysis," and, despite his reasonable objections to the term (56, p. 11), it still seems preferable to designate it as the "environmental focus." In the discussion of the principles of general systems, it will become apparent why "environment" is regarded here as the more fitting word.

Some phenomena bearing on or influencing international behavior represent passive or constant factors. These serve purposes something like raw materials from which international behavior is fashioned. At any given moment, there is a "potential"-material resources, brainpower, experience, influence, technology, geographical space and location, and other such resources-drawn upon in the carrying-on of international relations. There are limits to these "raw materials," so that action in international affairs is restricted by their relative availability and abundance. Napoleon could not draw heavily enough on his potential to win his Russian campaign. Between 1945 and 1949 the United States did not muster sufficient resources from its potential to prevent the Communist seizure of power in

China. In many instances, it can be established, after the fact, either that the actor did not perceive the limits of his available capability and, consequently, overcommitted himself in interaction or that the actor failed to recognize the existence of some important potential and suffered losses or disadvantages from oversight, misconception, or ignorance.

By recalling the two modes of explanation-culturally rooted and supracultural-we can now relate these two modes to the category of the "environmental focus." A division may be made between the theories, propositions, and empirical findings of the environment of the international system (which includes its actors and their interactions) according to whether they refer to a culturally mediated potential or to the potential wholly determined in the "works of nature" and therefore "supracultural." As an illustration of the culturally mediated potential, a government, having failed to organize sufficient military transport, may find that it cannot take the police action it desires to support some six thousand miles beyond its own territory. Limits are imposed on the action it would otherwise undertake. Various examples of limits arising from the state of nature run to the type of the impossibility for a man to be in two places at the same time or to the inability to build agriculture in a desert because of the lack of water.

Some conceptual clarification and, perhaps, a marked improvement in our theorizing would result if the practice of classifying international relations ideas in the terms of the specifications of the last few pages should become common. It was suggested above that a classification framework, intermediate between the subject names and the particular topics, should be built. For purposes of reference and inventory, we could place the theoretical mate-

rials already accumulated in this six-category structure:

actor focus culturally rooted mode	interaction fo- CUS culturally rooted mode	environmental focus culturally rooted mode
actor focus supracultural mode	interaction fo- CUS supracultural mode	environmental focus supracultural mode

No classification scheme sorts perfectly all the elements that are brought to it. The argument is only that the theoretical materials at hand for international relations are not sufficiently sorted and arranged by the customary classifications now existing for the field. The suggestion of the six-category scheme is offered in the hope that it may be useful in reducing some of the conceptual confusion now encountered and in increasing the utility of the bodies and elements of theory now in being.

Still greater gains would come to the discipline, not only in co-ordinating the present store of ideas but also in stimulating new formulations, if there were available a large neutral framework which would accommodate modes, foci, subjects, and topics along with numerous conceptual schemes and islands of theory. The attractiveness of the principles of "general systems" lies in this direction. It was a preoccupation with the broad problem of an emerging synthetic discipline which is obliged to borrow and adapt materials from many fields of knowledge that first kindled this writer's interest some half-dozen years ago in the possible applications of "general systems" to international relations.

Since then, others have also seen these possibilities, so that the term "system theory" has attained a degree of familiarity among the students of international relations. It has become evident, too, that mis-

understandings and misconceptions have grown, apparently because the breadth and variety of the possible applications have not been realized. Mainly, for the reason that many translations of general systems principles to the field constitute some good clues to the solutions of problems of theory construction in international relations, this paper concludes with a brief description of some of the ideas of general systems.

Contributions from the Systems Approach

The term "system" appears frequently in everyday discourse. An assertion that there is a "systems approach" to international relations would convey, therefore, some information to most individuals. It would appear to throw no new light on the problems of international relations study, however. The usual definition says that a "system" is an assembly of parts in a working whole. From this we might derive the unexciting observations that (a) nations (parts or entities) exist and (b) nations interact (parts operate in ensemble; entities take action and are acted upon). This strikes us with the force of truth-obvious truth.

If further support of the idea is needed, we may consider a statement about systems by Kenneth Boulding:

Whatever is not chaos, is system, and wherever there is system, there can be knowledge. Wherever there is subjective knowledge-what I have elsewhere called an image-there are expectations. . . . Where expectations are persistently *fulfilled*, it is reasonable to suppose that the image on which they are based is "true" in the sense that it represents knowledge of a system of which the knower is a part, or which the knower can observe [8, p. 1].

Most of us would not object to these thoughts; we believe that, no matter how complex they are, international relationships have a degree of orderliness and regularity

as well as a patterning in our apperceptions to the extent that our expectations are frequently fulfilled. We are not utterly and endlessly astonished or mystified by each new turn of the events of international affairs. There is some knowledge of the workings of the international system. What more is there to be said?

General systems principles are something more. Whenever phenomena are encountered which have unmistakable characteristics of complex organization, these principles become relevant. They are a *guide* to a form for the construction of the problem. The form develops in consideration of the following: We are aware that we face a situation of organized complexity, and we want to study and understand how the complexity "works." We have already found that any one factor, or influence, or trait isolated from the situation cannot be followed satisfactorily from beginning to end to account for what is going on because we see that this single element is a part of a crowd of other elements, flowing together, emerging, combining, recombining, changing, growing, and, in general, going through processes. We are in the position of being unable to determine what the single element does, relates to, and produces in the manifolds of combinations and interactions. There is little chance of maintaining a crisp separation of the parts. Furthermore, it is recognized that the single factor, influence, trait, or element is a simplification we have made in our minds, either conventionally or consciously, while, in fact, it is a complex organization itself, functioning in part in terms of its "inside" nature and in part according to the what, the where, and the when, of its associations and its interplay with other factors and elements. Finally comes the dawning awareness that the ordinary inspection of the organized complexity under study, the various intuitions, insights,

and hunches about how it operates, and the many judgments and evaluations about its nature and its effects all have resulted in some identification of the phenomenon as a whole but not in a satisfying explanation of the principles of its operations.

The students of international relations can decide for themselves whether or not our efforts have taken us about this far. The general systems approach is, in any case, a "strategy of the next step."

The first step is a conscious recognition that the general form of the problem is not unique. The fact that physicists, biologists, economists, and psychologists, among others, are dealing with problems of organized complexity is not the crucial consideration. It is rather that the struggles with the particular subject matters have yielded many ways of formulating plans of attack and that a number of features of these plans are of a common form. Let it be emphasized that it is the *form* alone that reappears in field after field.

That form is a primary mental construction which separates the phenomena being studied into two parts: a "system" and *its* "environment." There are three basic questions to be asked about "system" and "environment."

1. What are the operating parts of the system (these parts are usually called "components")-what do they do, how are they arranged, how are they co-ordinated, how do they fluctuate, change, or grow, and how are they replaced?

2. What are the boundaries between the system and environment-what functions do the boundaries serve, how are they structured, maintained, and changed?

3. What is the character of the influence of the environment on the system and of the system on the environment-how do these two hypothecated complexes interact?

The third question can be turned into the

first by a controlled reidentification of "the system." This is important and, among other things, carries us back to the earlier discussion of the "actor focus" and the "interaction focus." Let us take an "interaction focus" for a moment by thinking of an "international system" made up of the relationships of State A and State B. Under question No. 1, we may search for and find the constituting parts for both A and B which are the components of the system. We identify for A its parts called $a_1, a_2, a_3, \dots, a_n$ and for B the parts $b_1, b_2, b_3, \dots, b_n$. By doing this, we hope to find out about how and why A and B operate and perform the way they do *in the system*, and under question No. 2, how they are maintained within boundaries which separate their operations and performances from everything else in the world. In question No. 3, we should like to know, however, what in the world directly influences and is influenced by the workings of the system $A \leftrightarrow B$.

Now, shift the perspective and redefine the system: State A, now, *is* a system existing in the *international environment*-we know that this is a fact by experience and common sense. We now have an "actor focus," and question No. 1 applies: system A has components $a_1, a_2, a_3, \dots, a_n$ which, if we knew all about them, and precisely, would explain part of the workings of A. Question No. 2, when answered, will tell us just where system A "is," and question No. 3, under which we now find B along with some C's, D's, E's, \dots , directs us to the rest of what we need to know in order to tell how and why system A operates. In a different perspective, then, question No. 3 of the "actor focus" guides us to an examination, probably *less precise but more extensive*, of the materials appearing under question No. 1 of the "interaction focus." The same possibilities of redefinition exist for an "environmental" focus and, we can

see, in the strategy of general systems, a rationale and relationship for the three perspectives of actor, interaction, and environment which were treated in the preceding section of this paper.

Since the workers in many other fields also face the problems of organized complexity and also attack these problems with the same three basic questions related to system and environment, there is an opportunity to learn from them and to carry out, perhaps, some selective "horrorizing." Where the nature of the subject matter has permitted more thorough applications of the *form*, there may lie useful clues and leads.

We are not sufficiently armed with this conceptualization of the systems form, however, until we take into account certain other basic considerations. There are different orders of inquiry to be applied to the study of a complex system and its environment. The system, being complex, needs to be investigated from a number of different angles and in various ways. Some investigations must wait until others have been completed. An order of battle needs to be drawn up, in other words, and the different "orders of inquiry" are precisely the elements used in that kind of planning. Kenneth Boulding has provided a useful list of relevant "orders of inquiry:"

frameworks
clockworks
control mechanisms
self-maintaining structures
genetic-societal structures
image and "knowledge" structures
individual human-symbol using structures
social organization-symbolic structures
transcendental structures [7, pp. 14-16]

A list of this kind becomes significant only in the discussion which accompanies it. Not all of Boulding's levels will be considered here, since the immediate purpose

is only to illustrate the utility in recognizing and using such orders of inquiry.

Let us consider the order of questions related to "frameworks." In international relations we have long been preoccupied by framework questions and problems. We have sought to say what the state is, what nationalism is, and what specific patterns constitute diplomacy, imperialism, interests, power, international trade, conflict, the state system, co-operation, war, and peace. Knowing that all these "things" are variable, situational, and too complicated to be named in simple formulas, we have sought, nevertheless, to fix them as entities and constellations. Complaints against anthropomorphization, static analysis, and misplaced concreteness have not deterred us, and, indeed, our store of systematic knowledge is mostly concentrated in such descriptions of "frameworks." Identifying, locating, naming, and mapping are principal operations in framework inquiries. Boulding notes:

The accurate description of these frameworks is the beginning of organized theoretical knowledge in almost any field, for without accuracy in this description of static relationships no accurate functional or dynamic theory is possible. Thus the Copernican revolution was really the discovery of a new static framework for the solar system which permitted a simpler description of its dynamics.

(ii) The next level of systematic analysis is that of the simple dynamic system with predetermined, necessary motions. This might be called the level of *clockworks* [1, p. 14].

The order of our questions at this second level of "clockworks" shares, with many other fields, a central interest in the mechanisms of departure from and return to the state of equilibrium. What would we do in international relations without the explanations of the balance of power and the balance of trade? Without these and other quasi-equilibrium concepts, such as stimulus and response and the alternation of war

and peace in historical accounts, the study of international relations would lack, largely, a sense of movement and process. Yet the exploitation of more sophisticated types of equilibrium inquiry has not really begun in international relations. For example, we have little systematic work which places a number of fairly reliable and steady factors in a hypothetical equilibrium system and then intensively considers the probable effects of some "wild" factor from the environment on the working of the balance. With the change of level from "framework" to "clockwork," it is quite apparent that the type of questions changes also.

A change in the order of relevant questions occurs, no less, in the shift to the matter of control mechanisms. The ideas of communication theory, information theory, and cybernetics have been ignored or thrust aside by most students of international relations perhaps because the connection with many interesting questions in international relations has not been fully perceived. A veritable Hood of theoretical and research problems concerning international behavior springs forth when attention is directed to control properties of systems. For example, the international political system is recognized, even on the roughest level of approximation, as a very loosely articulated and integrated system. Is this an important control property in itself? A change or a stimulus in one part rarely travels through the whole structure. Yet, on occasion, a remarkably small and local "input" travels, amplifies, and reverberates through the entire structure. Two basic questions arise: What is the nature of the "guardians" of indifference (whatever they may be) which, most of the time, protects the stability and safety of the total system from disturbance and shock? This occurs in a way reminiscent of the control methods of a dynamite plant where each production and storage unit is

isolated from every other, reducing the danger of blowing up the entire site. There is, here, a passive control system which we know little about for international relations. On the other hand, the perturbation which spreads and grows proves that the system is not, *in its fundamental organization*, nearly as unintegrated as we usually suppose. This rare kind of disturbance may raise the level of irritability to a dangerous point. We may ask: What permits the amplification of a small disturbance? How does it break through the gatekeepers and guardians which, most often, are effective in maintaining general indifference or insensitivity? And what processes return the international system to its "normal" condition of apparent disjointedness?

There is a general proposition which states that, the more complexly organized a system is, the greater will be the need for more numbers and kinds of control and system maintenance (57). Does this proposition make suggestions and raise questions about trends toward increasingly complex international relations and the concomitant requirements for more international control organizations? Are there self-generating influences at work here? Is the United Nations a part of a feedback organization, and, if so, what types of transaction are corrected?

Beyond the level of control mechanisms, let us consider briefly only the next to the last of Boulding's categories—the symbolic awareness of social organization structure. Boulding's evaluation of the problem is interesting:

Beyond the second level [of clockworks] adequate theoretical models get scarcer. . . .

Beyond the fourth level [of self-maintaining structures] it may be doubted whether we have as yet even the rudiments of theoretical systems. . . . The kind of knowledge and skill that we have at the symbolic level is very different from that which we have at lower lev-

els-it is like, shall we say, the "knowhow" of the gene as compared with the knowhow of the biologist. Nevertheless it is a real kind of knowledge and it is the source of the creative achievements of man as artist, writer, architect, and composer.

Perhaps one of the most valuable uses of the above schemes is to prevent us from accepting as final a level of theoretical analysis which is below the level of the empirical world which we are investigating [7, pp. 16-17].

In the study of international relations, it is clear enough that the empirical world that we investigate is divisible in many useful ways into system, component, and environment, but always we are dealing with symbol-using human beings who are self-conscious and, in this age, aware also of their social organizations not only within sight and sound but also in the distance. Not only that; these symbol-using beings anticipate the actions of others and trim their behavior accordingly. When it is discovered that the other is anticipating and trimming, this, too, becomes a basis for further changes in behavior. The future is foreseen. An increased ability (say, from scientific accomplishment) to see into the future enters into the mixture to change not only expectations but the future as well. Social science aspires to understand all this and in such a way that its way of knowing can be shown in public. International relations has the responsibility to comprehend and explain a segment of the total social reality.

Social science has proceeded far enough with problems of frameworks and clockworks to have created at least a professional awareness of the need to use numerous complementary concepts such as society, culture, and personality. We know enough now to know we shall not answer the "why" questions by single-aspect or single-factor analyses. Social sums are different from the

sums of the social parts, and it is now commonplace to recognize the multifactor interactions in all social phenomena. In this connection it is encouraging to see that the study of international relations is not too far behind; some beginnings of international relations theory in this tier of complexity have been made, for example, in the recent inquiries into bargaining and strategy.¹⁵

In the perspective of general systems, international relations is seen as an organized complexity on a high human level, and, for understanding it, we have mostly the simplest of theoretical tools for the lower levels of complexity. In the systems focus, this theoretical underpinning can be regarded mainly as collections of *ad hoc* concepts and propositions whose value, in total, is considerable, nevertheless. The usefulness of this theory is not diminished by recognizing it for what it is. Its worth may be even enhanced by seeing it in terms of the orders of frameworks, clockworks, control systems, etc. It is a value of the general systems approach to help to clarify and rationalize the theory we already have.

Another function of general systems is found in its methodologies. There are two main types of methods for carrying on research in the systems framework. One might be called "microsynthesis" and the other, "macrosynthesis." About the first, Ross Ashby (3, pp. 4-5) has remarked that general systems offers "a single vocabulary and a single set of concepts suitable for the scientific treatment of the system in which complexity is outstanding and too important to be ignored." Ashby notes further:

¹⁵ Until lately, the literature on strategy and bargaining did not take international behavior and international situations within its focus, except rarely. The Change is illustrated, for example, by the number of articles bearing on these subjects and appearing in this *Journal* in recent months (47, 48, 39, 15, 52).

The emergence of general system theory is symptomatic of a new movement that has been developing in science during the past decade: Science is at last giving serious attention to systems that are intrinsically complex. This statement may seem somewhat surprising. Are not chemical molecules complex? Is not the living organism complex? And has not science studied them from its earliest days? Let me explain what I mean.

Science has, of course, long been interested in the living organism; but for two hundred years it has tried primarily to find within the organism, whatever is simple. . . . From the whole complexity of digestion, the biochemist distinguished the action of pepsin on protein, which could be studied in isolation. And avoiding the whole complexity of cerebral action, Pavlov investigated the salivary conditioned reflex—an essentially simple function, only a fragment of the whole, that could be studied in isolation.

The same strategy of looking for the simple part has been used incessantly in physics and chemistry. Their triumphs have been chiefly those of identifying the *units* out of which the complex structures are made. The triumph has been in analysis, not in synthesis. . . .

Thus until recently, the strategy of the sciences has been largely that of analysis. The units have been found, their properties studied, and then, somewhat as an after-thought, some attempt has been made to study them in combined action. But this study of synthesis has often made little progress . . . [2, p. 1].

Ashby's method is a procedure for following the small discrete steps of fluctuation or change which occur within a well-defined action system or in the exchanges of materials between system and environment. The successions of events and their "transformations" which are too numerous and varied to keep in mind by normal means are, in effect, traced and tracked through their circuits by the use of information theory or cybernetics. This is not analysis but synthesis, and, if it should be applied in international relations, it would probably deserve a "microsynthesis" label. An application may be possible in the de-

tailed explorations of problems such as those indicated in Snyder's conceptual scheme for decision-making.

Macrosynthesis, in terms of general systems, would appear to apply to international relations in a "try-for-fit" procedure for large numbers of properties and operating characteristics of open systems that have been noted to appear in many different fields and phenomena. We return, here, to the "big-slab" approach to theory. It would be worthwhile to have an inventory or listing of propositions about complex open systems that might be tried against the data of international relations. A few illustrations will be given.

"Power" may be defined in systems terms as the efficiency with which a system is able to give and get from its environment, whatever the need, purpose, or goal may be. As we know well, many factors play a part in making power, but one of the manipulatable factors is the elaboration of organization. Increasing organizational complexity can create phenomenal power increases under "right" conditions. A step-up in power through organizational elaboration usually involves, however, other effects, among them, increased vulnerability and increased conservatism. The abilities to survive disruptions and to change with environmental changes ordinarily are reduced. There are, of course, ways to offset these liabilities to some extent. The question raised is this: We have studied international politics intensively in terms of power but scarcely at all in terms of vulnerability and conservatism. Are not these comparably important to study and explain?

A major property of systems is "boundary maintenance." Biologists have paid careful attention to the types, functions, and performances of membranes, skins, shells, and other separators, liners, covers, protectors, and exchangers of systems interacting with

environments. The wide range of functions is impressive. In international relations the importance of territoriality is recognized and is a subject of law. Disputes over geographical frontiers have entered the subject matter frequently, but theorists have given scant attention to the topic (60). But the components and systems of international relations have many boundaries other than lines on territory. The system of action of United States foreign relations does not remain fixed; it expands and contracts even during relatively short time periods and boundaries must move in and out, changing their shape and, perhaps, their nature and functions accordingly. What are these boundaries like? Further, what are the boundaries of the United States oil industry considered in a world-wide system of oil supply, distribution, and consumption? There are many unexplored topics having to do with boundaries in international relations. John Herz has provided a major theoretical addition where we had virtually no theory before through his introduction of the concept of the changing permeability of the "hard shells" of national territoriality (18).

A final example of "tests for fit" for borrowed and translated systems concepts and methods is taken from a sector of theory and research which presently is very active in the study of international relations. This is the subject concerning community-building and integration. It says nothing against the interesting work that has been published lately to observe that fertile proposals and successful techniques of study in other fields are related to this current interest and, yet, have been mostly passed over. We have not carried into the field of international relations such typifications as Werner Landecker's four categories of integration (29), such general concepts as Boulding sees in the principle of nucleation (7, pp. 70-71), and

such measurement techniques as in Robert Angell's studies of the social integration of cities (1). There are a great many more possibilities for theory and empirical research in these and other regions of systems concepts and phenomena.

It will be objected that we do not need the trappings of any general systems approach to find problems to investigate in international relations. That is true, in principle. Yet one has no easy way to think about a subject which has not yet come to mind. The checklist or inventory of potentially applicable general systems propositions is probably extensive, and, who knows before the fact, what useful concept, point of view, approach, procedure, or lead might be found by turning to the literature of systems?

A fifth useful function of the general systems approach might be regarded as normative. Ludwig von Bertalanffy (4), among others, has demonstrated how biological as well as psychological circumstances pinch down the scope of our perceptions of reality to little more than a small viewing window. What is true depends in some part on the angle from which we view it and in some further part on how we manage to extend the range of vision from the window. In investigating organized complexity, theoretically or empirically, we need not so much merely to tolerate different ways of viewing as to encourage actively the discovery of different avenues of access to the phenomena. It is sometimes said that we already know all that we need to know about international relations or that we have more theory than we can use. Theorizing efforts from the systems standpoint can scarcely be called "a huge misstep in the right direction" (21, p. 40) if they widen at all the view of the phenomena or set forth additional and different ways of studying the data of international relations. The longer

one considers general systems in their full scope, ranging across various kinds of phenomena, the more lively becomes one's realization of how perilously small and uncertain is the state of really demonstrable knowledge. The person who is interested mainly in the social sciences not only learns how really modest his mien should be but also gains the valuable realization that there is no basis for feelings of deference toward the "magicians" of contemporary science—the physical scientists. The latter are, indeed, highly competent, but the problems they have solved by hard intensive work and by intelligence are, nevertheless, simple problems compared with the problems of the rest of science. The systems approach is one of the better antidotes against misapplications of methods carried from one field to another.

Further, the general systems perspective reveals clearly that, as the subject matter ranges from the physical sciences, through biology, to the social sciences, the problems of complexity increase radically, while the tools of method and of theory change from excellent in the physical sciences to fair in biology and to seriously inadequate in the social sciences.

There is no longer any doubt about the immense need for a sound and systematic knowledge of international behavior, international processes, international capabilities, and international institutions. There has even been, in recent months, a call for the establishment of a government-sponsored laboratory for a "science of peace" (51, 62). The students of international relations, as a professional group, probably know more today about international phenomena than any men in history, but it still is not enough. Two defects that are partially remediable stand out: (a) the accumulated knowledge, as great as it is, runs in only a few of the many channels of the "reality" and (b) the

variations in its quality and utility have not been adequately differentiated. We have not as yet sorted the knowledge that is demonstrable or could be made demonstrable from that which is not. Theory should be regarded as a means toward the improvement of demonstrable knowledge. A place for theory in the study of international relations is already assured; its role seems certain to gain and grow.

REFERENCES

1. ANCELL, ROBERT C. "The Social Integration of Selected American Cities," *American Journal of Sociology*, XLVII (1942), 575-92.
2. ASHBY, W. ROSS. "General Systems Theory as a New Discipline," *General Systems Yearbook*, III (1958), 1-17.
3. ——. *An Introduction to Cybernetics*. New York: John Wiley & Sons, 1956.
4. BERTALANFFY, LUDWIG von. "An Essay on the Relativity of Categories," *Philosophy of Science*, XII (1955), 243-63.
5. BINDER, LEONARD. "The Middle East as a Subordinate International System," *World Politics*, X (1958), 408-29.
6. BOCK, KENNETH E. *The Acceptance of Histories*. Berkeley: University of California Press, 1956.
7. BOLDING, KENNETH E. "General Systems Theory—the Skeleton of Science," *General Systems Yearbook*, I (1956), 11-17.
8. ——. "Political Implications of General Systems Research." Presidential Address, December 27, 1957, Society for General Systems Research. 19pp. (Mimeographed.)
9. BRECHT, ARNOLD. *Political Theory*. Princeton, N.J.: Princeton University Press, 1959.
10. CLAUDE, INIS L., JR. *Swords into Plowshares: The Problem and Progress of International Organization*. New York: Random House, 1956.
11. DAHL, ROBERT A. "The Concept of Power," *Behavioral Science*, II (1957), 201-15.
12. EASTON, DAVID. "Harold Lasswell: Policy Scientist for a Democratic Society," *Journal of Politics*, XII (1950), 450-77.
13. FERNBACH, ALFRED. "The Doctorate in International Relations," *Journal of Politics*, XVII (1955), 275-86.

14. Fox, WILLIAM T. R. (ed.). *Theoretical Aspects of International Relations*. Notre Dame, Ind.: University of Notre Dame Press, 1959.
15. GLADSTONE, ARTHUR. "The Conception of the Enemy," *Conflict Resolution*, III (1959), 132-37.
16. GuETZKOW, HAROLD. "Long Range Research in International Relations," *American Perspective*, IV (1950), 421-27.
17. HAAS, ERNST B. "Persistent Themes in Atlantic and European Unity," *World Politics*, X (1958), 614-28.
18. HERZ, JOHN H. *International Politics in the Atomic Age*. New York: Columbia University Press, 1959.
19. ——. *Political Realism and Political Idealism*. Chicago: University of Chicago Press, 1951.
20. HoBSON, J. A. *Imperialism: A Study*. Rev. ed. London: Allen & Unwin, 1938.
21. HOFFMANN, STANLEY (ed.). *Contemporary Theory in International Relations*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1960.
22. INTERUNIVERSITY SUMMER SEMINAR ON POLITICAL BEHAVIOR. "Research in Political Behavior," *American Political Science Review*, XLVI (1952), 1003-45.
23. JENKIN, THOMAS P. *The Study of Political Theory*. Garden City, N.Y.: Doubleday & Co., 1955.
24. KAPLAN, MORTON A. *System and Process in International Politics*. New York: John Wiley & Sons, 1957.
25. KAUFMANN, WILLIAM W. "The Crisis in Military Affairs," *World Politics*, X (1958), 579-603.
26. KRSSINGER, HENRY A. "The Conservative Dilemma: Reflections on the Political Thought of Metternich," *American Political Science Review*, XLVIII (1954), 1017-30.
27. ——. *Nuclear Weapons and Foreign Policy*. New York: Harper & Bros., 1957.
28. KNORR, KLAUS E. *British Colonial Theories, 1570-1850*. Toronto: Toronto University Press, 1944.
29. LANDECKER, WERNER S. "Types of Integration and Their Measurement," *American Journal of Sociology*, LVI (1951), 332-40.
30. LANGER, WILLIAM L. "A Critique of Imperialism," *Foreign Affairs*, XIV (1935), 102-19.
31. LASSWELL, HAROLD D. "The Political Science of Science," *American Political Science Review*, L (1956), 961-79.
32. LERCHE, CHARLES O., JR., and SAPIN, BURTON M. (eds.). *Some Problems in the Study and Teaching of International Politics*. Atlanta, Ga.: Emory University Press, 1958.
33. LERNER, DANIEL. *The Passing of Traditional Society: Modernizing The Middle East*. Glencoe, Ill.: Free Press, 1958.
34. LUNDBERG, GEORGE. "Semantics and the Value Problem," *Social Forces*, XXVII (1948), 114-17.
35. McCLELLAND, CHARLES A. "A Classification of International Relations Theory," *Prod*, II (1959), 32-34.
36. ——. "Systems and History in International Relations: Some Perspectives for Empirical Research and Theory," *General Systems Yearbook*, III (1958), 221-47.
37. McDoUGAL, MYRES, S., and LASSWELL, HAROLD D. "The Identification and Appraisal of Diverse Systems of Public Order," *American Journal of International Law*, LIII (1959), 1-30.
38. MANNING, C. A. W. *The University Teaching of Social Sciences: International Relations*. Paris: UNESCO, 1954.
39. MILBURN, THOMAS W. "What Constitutes Effective Deterrence?" *Conflict Resolution*, III (1959), 138-45.
40. MoRGENTHAU, HANS J. "Area Studies and the Study of International Relations," *International Social Science Bulletin*, IV (1952), 1-10.
41. ——. *Politics among Nations*. 2d ed. New York: Alfred A. Knopf, 1956.
42. ——. *Scientific Man vs. Power Politics*. Chicago: University of Chicago Press, 1946.
43. PADELFORD, NORMAN J., and LINCOLN, GEORGE A. *International Politics*. New York: Macmillan Co., 1954.
44. PARSONS, TALCOTT, and SHILS, EDWARD A. (eds.). *Toward a General Theory of Action*. Cambridge, Mass.: Harvard University Press, 1954.
45. PLATT, JOHN RADER. "Style in Science," *Harvard Theological Review*, CCXIII (1956), 69-75.
46. RUSSELL, FRANK M. *Theories of International Relations*. New York: Appleton-Century-Crofts, 1936.
47. SCHELLING, THOMAS C. "Bargaining, Communication, and Limited War," *Conflict Resolution*, I (1957), 19-36.

48. SCHELLING, THOMAS C. "The Strategy of Conflict: Prospectus for a Reorientation of Game Theory," *ibid.*, II (1958), 203-64.
49. SCHWARZENBERGER, GEORG. *Power Politics*. New York: Frederick A. Praeger, 1951.
50. *Science*, CXXIII (1956), 93.
51. "Science in the News," *Science*, CXXX (1959), 966-67, and CXXXI (1960), 340.
52. SINGER, J. DAVID. "Threat-Perception and the Armament-Tension Dilemma," *Conflict Resolution*, II (1958), 90-105.
53. SNYDER, RICHARD C., BRUCK, H. W., and SAPIN, BURTON. *Decision-making as an Approach to the Study of International Politics*. Princeton, N.J.: Foreign Policy Analysis Project, 1954.
54. SONDERMANN, FRED A. "The Study of International Relations: 1956 Version," *World Politics*, X (1957), 102-11.
55. SPROUT, HAROLD. "International Politics and the Scholar," *Princeton Alumni Weekly*, LIX (1959), 8-10 and 14-15.
56. SPROUT, HAROLD and MARGARET. *Man-Milieu Relationship Hypotheses in the Context of International Politics*. Princeton, N.J.: Center of International Studies, 1956.
57. TERRIEN, FREDERIC W., and MILLS, DONALD L. "The Effect of Changing Size upon the Internal Structure of Organizations," *American Sociological Review*, XX (1955), 11-13.
58. THOMPSON, KENNETH W. "The Study of International Politics: A Survey of Trends and Developments," *Review of Politics*, XIV (1952), 433-67.
59. ---. "Toward a Theory of International Politics," *American Political Science Review*, XLIX (1955), 733-46.
60. WHITTLESEY, DERWENT, *The Earth and the State*. New York: Henry Holt & Co., 1944.
61. VOLFERS, ARNOLD, and MARTIN, LAURENCE W. (eds.). *The Anglo-American Tradition in Foreign Affairs*. New Haven, Conn.: Yale University Press, 1956.
62. WMCHESTER, QUINCY. "Project for a World Intelligence Center," *Conflict Resolution*, I (1957), 93-97.
63. ---. *The Study of International Relations*. New York: Appleton-Century-Crofts, 1955.
64. ---. *A Study of War*. 2 vols. Chicago: University of Chicago Press, 1942.
65. ---. "The Value for Conflict Resolution of a General Discipline of International Relations," *Conflict Resolution*, I (1957), 3-8.
66. ZIMMERN, SIR ALFRED (ed.). *University Teaching of International Relations*. Paris: International Institute of Intellectual Cooperation, 1939.

On the Role of Metatheory in the Academic Discipline of International Relations

Submitted by Lucas Grassi Freire to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Politics
In September 2012

#his thesis is available for Library use on the understanding that it is copyright material and that no \$otation from the thesis may be published %ithout proper ac&no%ldgement'

I certify that all material in this thesis which is not my own or has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University'

Signature()()()()()()()()()()()()()()()()()()()''

On the Role of Metatheory in the Academic Discipline of International Relations

This thesis investigates in three parts the role played by metatheory in the discipline of International Relations. Part one defines metatheory as 'systematic discourse about theory' and classifies it in a typology combining elements internal or external to the discipline 'with intellectual or contextual aspects of theorising'. Each combination has particular functions. They also add to the roles played by several modes of metatheoretical inquiry: hermeneutical, evaluative, corrective, critical and historical. The typology offered in part one clarifies the general roles of metatheory as a constraining and enabling discursive mechanism. This is also discussed in part two addressing how scholars portray metatheory's role in the discipline. Arguments against and in favour of metatheory are scrutinised, leading to a qualified defence of metatheoretical research in IR. Some of the negative impact of metatheorising in IR is acknowledged, but ultimately a stronger case attempting to eliminate it from the field cannot be sustained for analytical reasons. The merits of metatheory therefore will depend on how it operates in particular instances. A selection of illustration cases in part three further develops the argument. The first case stresses how metatheoretical directives shaped 19th century views of the Ottoman Empire. It indicates that metatheory can frame theoretical claims even in a weak disciplinary context. A stronger disciplinary environment frames the second case, analysing a number of IR theories on the impact of the Peace of Westphalia in the European state system. This discussion often alludes to the notion of hierarchy. The third case examines the interaction between metatheoretical directives and theories of hierarchy. These arguments are not necessarily compatible with the metatheoretical principles argued by their authors. As a mechanism therefore, metatheory does not relate to theory in a deterministic way. Part three itself is of course a metatheoretical study that further illustrates the thesis.

Contents

Introduction	7
Metatheory as an issue in IR	7
IR studies on aspects of metatheory	""
The argument under consideration	"9
 PART ONE	
 1. Theory in the Discipline of International Relations	"
Introduction	"
Context in philosophy and social science	6
Theory in classical, behavioural and normative IR	3
Science and theory in mainstream IR	6
IR theory as social theory	63
IR theory as discourse and practice	6
Questions on theory and IR	:<
Another level of discourse	=:
 What is Metatheory?	=9
Theory of theory	=3
Metatheory in cognate disciplines	=7
Typology of metatheorising	9
Metatheoretical research	93
Synthesis and clarifications	3"
Final remarks	3=
 PART TWO	
 2. The Nature & Uses of Metatheory in International Relations	33
Introduction	33
Intrinsic features of metatheory	3<
The complexity of metatheory	<=
The teaching of metatheory	7
The politics of metatheory	73
Discussion	"!6
Final remarks	"!<
 3. The Position & Uses of Metatheory in International Relations	""!
Introduction	""!
+ tool for improving IR theory	"""
+ tool for understanding the IR discipline	""=
+ tool for understanding theoretical material	""<
+ tool for understanding theories in the social world	"
Discussion	" 3
Interlude	"6"

PART THREE

*. Metatheory and International Political Theory+	"66
The Holy Roman Empire in Early Modern thought	
Introduction	"66
#he Empire as universal public association	"69
#he Empire as a monstrous political body	"."
#he Empire as a union and state	"=
Metatheory in international political theory	"7
Final remarks	"=3
.. Metatheory and Theoretically-Oriented History in IR+	"9"
Narratives on the Peace of Westphalia	
Introduction	"9"
Westphalia2 legitimate authority and anti-hegemony	"9
Westphalia as the constitutional order of a modern system	"97
Westphalia as consolidation of a pre-modern order	"39
Discussion	"<"
Final remarks	"<<
.. Metatheory and a Theoretical Notion+	"7"
Hierarchy in World Politics	
Introduction	"7"
+ note on the context of the hierarchy literature	"76
*ierarchy and weaker states in neorealism	"7=
Mainstream theorising and dyadic hierarchy	"7<
Normative theory and hierarchy in peripheral realism	!
Deductive theory2 time-preferences2 wealth and power	!3
Eurocentrism and hierarchy in international thought	"
Final remarks(levels of metatheoretical analysis	"<
Conclusion	
+ summary of the primary contribution	
Secondary contributions	6
Questions2 answers and evaluation	=
Final remarks	3
Bibliography	7

Tables and Figures

Table 1.1 ? @ie% of theory in 1-	:75=!
Table 1.2 ? 8etatheorising according to focus2 %ith examples	9:
Table 1.3 ? 8etatheorising continuum2 according to focus	99
Table 1.4 ? .ombined foci2 relations and %ays of metatheorising	3
Table 1.5 ? >egative claims against metatheory classified	" 7
Table 1.6 ? Positive claims for metatheory classified	"6!
Figure 1.1 ? Politics as symbiotic consociation	"6<
Figure 1.2 ? #he universal public association	" :!
Table 1.7 ? #hree 1- accounts of the Peace of 4estphalia	"< 56
Table 1.8 ? #heory5history discussions in the typology of metatheory	"<6
Table 1.9 ? -oles of the metatheoretical study of discourses on 4estphalia	"<=

Acknowledgements

Many thanks to

My parents and sister for their love, care, support and positive influence. I dedicate this thesis to them. My family, especially Gonno, Ana and my cousins. Friends, in particular Leonardo Ramos and Ben P. Jett, who provided helpful suggestions, encouragement and incentive.

Special thanks to Colin Wright, an outstanding teacher, supervisor and mentor. Thanks to Sandro Eovisto, Lina Ampsher and Joice Aiguashca for their support as part of the supervisory team.

Exeter University for fully funding my 8+ programme with the International Full-Fee Scholarship awarded in 2013, as well as the doctoral programme via the Overseas Research Scheme. The Department of Politics for offering a bursary to cover living expenses in 2013. Proton Products for providing conference attendance support with the Global Awareness Conference awarded in 2014. The International Studies Association for partly funding my trip to Porto so I could present some of the partial findings of this thesis in 2014. The Institute for Humane Studies for offering a place at their Scholarship and a Free Society programme in 2014.

Bonathan Sunday, Eleni and Bulis Anspornsen and Geraldo Silveira for being excellent role models and for their pastoral care. David Armstrong, Eduardo Gouveias Silva, Gonfre Santos Filho, Paulo Luis Esteves, Katilde de Souza, William Ricardo SH, Baviera Adell and other teachers in Law and Economics whose support was crucial to my early intellectual growth. Eugenio Dini and Hugo E. Gama Pereira, early supervisors and mentors. Daissin Lopes, Guilherme Soares, Erin Javier, who led the way from PU Minas to Law scholarship. Liana Lopes, who predicted I would obtain a PhD abroad and Silvana Eurico, who told me about Exeter University.

My Law editorial team for proving the hypothesis that it is possible to write about world politics and have fun at the same time. Claudio Adelli for refining the

hypothesis by adding 'political engagement' to 'writing' and 'having fun'. Bim Sillen, Bonathan, Japhlin, Guilherme, @arvalho, David Eoy, Cis, Bob Boustra, Omer, Agares and Simon Polinder, scholars who encouraged me to reflect more on the role of philosophical assumptions in my discipline. Heofanis, Exada, Tylos, Af, Fatani, Samuele Dossi, Beremy, Aildeman for sharing on top of friendship and support innovative research tools and the latest news about academia in the UK and Europe. Bac, Sue, Fox, Chris Longman, Sarmita Ghosh, Susan Targetts, +lastair, Roc & Er for support and encouragement since my 8+ period in Exeter.

Jice, Aiguashca, Bohn, *eathersha, Jim Dunne for their intellectual and practical support with my research and academic life and for welcoming me to the Department of Politics as a student and teaching assistant. Ornelia, Avari, Igel, +shford and Bohn Elliott for mentoring me with an impact perhaps much larger than they imagine. Lui, C. Feldman for sharing some of his practitioner's insights on how theoretical issues apply or do not apply to the real world.

Ansour, Sasra for offering the opportunity to redesign and convene the 'Contemporary Theories of World Politics' module together. Andy Schaap for supporting the idea. Iain, *ampsher, Son & for all the assistance provided in the final months and during the research methods training. Nicole, Olleyer for helping me clarify the research topic. Sar, Do, Eovisto, *elen, Hurton and Gwen, Thomas for offering the opportunity to work together as co-authors and workshop organisers.

Exeter's Political Theory Reading Group, especially Sarios, Filis, Chris Fear, Robin Douglass, Simon, *onsend, Jiao, Khang and Stuart Ingham for their comments on an embryonic version of 'Chapter 1'. Professor Sarco, *uesbe, Llanos for supplying some of the secondary literature. Daniel Green and Huran, Eayaoglu for their comments on the '1st IS+ paper conference based on the same chapter'. Exeter Postgraduate Research Seminar members for asking relevant questions on my presentations at several stages of the research, in particular my office mate GC & Er, Eocadal, Audrey, +leDandro, +ndre, Lin & later and Patric & Haddeus, Bac & son for their comments on the '1st 4IS conference and Exeter Postgraduate Workshop paper introducing some of my critical evaluation of the 1st literature on metatheory'.

Of course I am responsible for this work including possible mistakes and eventual imprecision. *o%ever, ac & no%ludging their interaction with it does justice to my perception that I was welcomed and encouraged to express my views and subjected them to criticism.

Special thanks to Emma Elliott for her friendship, support, care, encouragement,
incentive and love'

In Ecclesiastes /" "' +@02 Eing Solomon %isely %arns me(

+nd further, by these, my son, be admonished,
of making many books, there is no end,
and much study is a weariness of the flesh'

Over the years, the presence and support of the wonderful people mentioned here have
constantly reminded me that MwearinessN can be avoided if Mmuch studyN is understood
as one among many facets of life ? a great one, to be sure, but not its chief end'

L'G'F'

Exeter, September 2012

Introduction

Metatheory as an issue in IR

In this thesis I investigate the role played by metatheory in the academic discipline of International Relations (IR). As a provisory definition, I start with the notion of metatheory as 'theory of theory' or 'systematic discourse on theory'. imilar designations can be made, such as 'scholarship about scholarship', as a kind of 'bridge' leading us beyond theory, or perhaps connecting the generalities of philosophical inquiry to the specificity of research in a given field. Arguably, every academic discipline is potentially or actually engaged in discussions about its own nature as a field of knowledge. Examples abound: philosophy of mathematics, of physics, of biology, and the philosophy of science. In the social sciences, economic 'methodology', social metatheorising, social philosophy, and the philosophy of social science. In humanities, philosophy of history, metaethics, prolegomena to theology, and philosophy. It comes as no surprise that a similar discursive domain has gradually emerged in our discipline since IR became institutionalised with its own self-identified specialists, theories and canonical texts. Yet, most of this discussion concentrates on localised aspects of metatheorising. Important as they are, these studies lack clarification on what this discursive layer does for IR as a discipline in general terms.

Unlike the self-reflective sub-disciplines mentioned above, there seems to be no clear identification of 'metatheoretical IR' as a sub-field yet, although instances of 'debates' at the level of metatheoretical discourse have been identified. There are several possible explanations for this fact. One is that IR is a relatively young academic discipline and shall eventually circumscribe metatheory as a domain of its own. Another position is that there still is much confusion about the nature of theoretical inquiry in the field as such, and what is often metatheoretical seems to be portrayed as primarily theoretical. Finally, others could possibly ascribe this state of affairs to the general situation of confusion on what IR is supposed to study in the first place. Having prevented the establishment of a widely recognised sub-field of metatheory in the discipline, these three factors taken together would apparently discourage my endeavour here. Instead, I take them rather as stimulating challenges and motivation.

On the first view that IR is a young discipline I grant that this is indeed the case, but the assumption that IR will inevitably obtain its own 'metatheoretical sub'field' rather than remain confused on this matter may seem too optimistic for those acquainted with the nature of metatheoretical scholarship in the discipline. There is no necessary reason why metatheory in IR would clarify itself if only we gave it more time, but it is indeed my hope that this thesis will work to that end by defining what metatheory is, and the main roles it plays in shaping academic inquiry in IR. This leads me to the second view on the status of metatheoretical discourse in the discipline namely, that IR scholars barely define 'theory', let alone 'metatheory'. By defining 'metatheory' and deriving its implied general roles, I provide a contribution in this respect. Because I start from the generally accepted notion of metatheory as 'theory of theory', there is a sense in which the meaning of 'theory' in IR itself is assessed here. Granted, there is no disciplinary agreement on what IR theory is (or should be) and, by implication, what it aims to do, how we test it, how we tell its story, and so on (3. & 4. Jackson, 5/..a). However, this should not necessarily prevent us from further developing metatheoretical discourse in IR. As we shall see, the same word, 'theory', is employed with different meanings in the discipline, but it is possible to distinguish how each group of theorists employs it and where they converge. Collateral issues like the goal of theory, its use, test, history, etc. can only be assessed if we engage in 'systematic discourse on theory', or simply, 'metatheory'. That is far from being a reason to avoid metatheoretical research, the disagreements on the role of theory in IR are rather an invitation to metatheorise and, therefore, reinforce the need for something like the present thesis. Finally, the third view on general confusion on what IR is meant to study, and whether it should metatheorise at all once again invites us to this exercise. For, although the question of whether IR should be studying the globalisation of sports or the transnationalisation of religion pertains to another level, it is surely the case that forming a general self-image of what the discipline is or should be pertains to the nature of metatheory as 'scholarship about scholarship'.

That is the role of metatheorising in general⁸ What roles does it play in IR⁸ In what ways does metatheory shape theoretical research in the discipline⁸ In the present thesis I address these questions by developing and illustrating in concrete terms the following argument% metatheory plays roles in IR that relate both to (.) the sub'ject' matter addressed by metatheoretical discourse and (5) to the mode of metatheoretical inquiry. On the sub'ject' matter side, we may divide metatheory in terms of combining its

focus on internal or external to the discipline of IR with its 'point of entry' into theory in intellectual or contextual terms. Thus, one way of identifying what metatheoretical discourse does is by looking at these permutations: Internal: Intellectual, Internal: Contextual, External: Intellectual and External: Contextual. As for the modes of metatheoretical inquiry, I analyse four main 'ways of metatheorising': hermeneutical, evaluative, corrective, critical and historical. Metatheory, following this scheme, (a) interprets theory; (b) judges theory according to certain standards; (c) refines theory; (d) provides social critique of theory; (e) accounts for the formation of theory over time. These are the main roles of metatheory at the most general level. In the first part of the thesis, I explain each of them in a conceptual manner. In the second part, I look at what IR scholars themselves consider to be the effect of metatheorising in the discipline.

Some have a very negative opinion of it, others understand it can play a positive role. There is a sense in which the IR literature goes beyond the conceptual generalities addressed in the first part of the thesis. The second part, in this vein, draws on what IR scholars have written in order to include an account of those roles played by metatheory which are contingent upon the situation in the field. Another way to make the discussion palpable is to illustrate the operation of metatheory in a number of 'illustration cases'. In the third part of the thesis, therefore, I look at the several ways in which metatheory shapes theoretical inquiry in a selection of texts and debates in IR and some of its sub-fields.

IR studies on aspects of metatheory

The relevant metatheoretical literature in IR can be organised in at least five types of text, examined below and in the following chapters. The first comprises systematic discussions on 'theory' by key theorists in the discipline. The second type is similar, but more focused on methodological texts with metatheoretical passages. These first two types of text pertain to a group of works which most scholars in the discipline would almost immediately recognise, given their influence in shaping the field one way or another. The remaining types are recent studies focused on specific aspects of metatheoretical research. The third category indicates the growth of this intermediate 'discursive layer' in IR with books specialised in the application of philosophy to our discipline. The fourth category involves not only 'disciplinary history', but also studies of 'theory-in-context' (that is, the impact of theory on social context and disciplinary dynamics, and vice versa). The final type of metatheoretical research relevant to my

argument includes a couple of recent journal debates on particular elements of metatheory in IR. Morgensen's (5/.) textbook containing advice on how to theorise world politics is a rare attempt to clarify some of the core issues in the field to the student audience, and an indication of a possible direction that the forthcoming literature might take.

Many IR scholars have on occasion discussed the nature of theory and scholarship. Classical theorists like Raymond Aron (1954, pp. 19, 20; see Frost, 1992) and Morgenthau (see Morgenthau, 1948, 1954) went as far as to address the role of IR as a discipline in light of the increasing popularity of positivist philosophy in politics and the social sciences. Kenneth Waltz's introductory chapters in *Theory of International Politics* (1979) on the character of theory and how one should theorise world politics from a 'systemic' perspective figure among the most controversial pages in the discipline (Joseph, 5/). This is attested by a range of critical and fertile responses to his approach (e.g. Ashley, 1988; Mahoney, 1988; Wendt, 1989). New theories emerging from this reaction in dialogue had to find their own discursive space in tension with mainstream scholarship under the influence of Waltz and others. The new disciplinary space was opened up by a series of metatheoretical pronouncements against mainstream theorising, its notions of social science and normative implications (e.g. Adler, 1992; Ceriani D'Amico, 1989; Eapen, 1989; Waller, 1996). Robert Mahoney's (1988) reinforcement of the dominant IR position on these issues triggered similar reactions (e.g. Weber, 1996). Scholars operating in alternative research programmes in IR often establish metatheoretical credentials before introducing or further developing theory (e.g. Wicner, 1995; Wendt, 1999). Methodology texts also reflect this type of tension, even though the metatheoretical issue is subordinated to the broader debate on research methods. Popular not only among IR scholars but also in general political science, some of the mainstream texts claim to provide a universalistic view of scientific research which has to be adopted in these fields in order to determine their robustness (King, Mahoney, D'Herba, 1996; Hanf, 1992). Critical reactions often emphasise theoretical and metatheoretical plurality and the desirability of reflecting this plurality in the choice of methods, as well as in the evaluation of empirical research (Eichbach, 5/; D'Amico, 5/). These debates on what makes good IR scholarship are often seen by theorists themselves as a source of division in the discipline (e.g. George, 1989; Holsti, 1989; Jones, 5/5), although many also locate the source in normative presuppositions

embedded in each research programme (e.g. 60\$, 5//@; Beufeld, .99A; 7alt, .99@). While displaying a metatheoretical character, these writings have not provided any extensive commentary on the general role of metatheory in IR.

A second group of texts in the metatheoretical literature in IR comprises the remaining types of studies focused on particular aspects of metatheoretical research. The first kind of study in this group deals specifically with applications of philosophical debates to the discipline. Martin Hollis and Steve Smith's *Explaining and Understanding International Relations* (.99/) is an earlier attempt to map IR along metatheoretical lines that helped to advance the agenda of disciplinary diversity. A number of discussions was triggered by that book (Hollis & Smith, .99.; Abner Cohen, .99=; Wendt, .99.), which provided a platform and a disciplinary 'vocabulary' for subsequent epistemological debates (George, .99G; Bicholson, .99=a). A key issue emerging from these earlier debates had to do with the implications of 'positivism' in IR (portrayed as the mainstream position on how the discipline should proceed) and how to react to it. Alexander Wendt's (.9@>, .995, .99@) work and, later on, Colin Wright's (.99=, .999, 5//5) contribution pointed out the need to look first at the diversity in ontological presuppositions, rather than the epistemological issue.

In the current metatheoretical debates in IR, epistemology and ontology frequently provide a link to methodological discussions and other elements of research. Particular positions in the philosophy of science are frequently 'imported' and adapted to our discipline (e.g. Herman & Herman, 5//5, 5//F; Joseph, 5//>; Joseph & Wright, 5//.). Fred Mervin's *Theory and Metatheory in International Relations*, in fact, defines metatheory in IR in this narrow way, i.e., as an application of the philosophy of science and of social science to our discipline (Mervin, 5//>, pp.=@ff). His own contribution purports to defend a conventionalist position and illustrate how this point of view would deal with issues of theory choice influencing policymaking (Mervin, 5//9b). In close dialogue with both philosophy and general social science, Wright's application of scientific realism adds to Wendt's focus on ontology and further advances a (philosophically) realist perspective on social agents and structures (Thomas Wright, 5///; 6. Wright, 5//=). (19a <url>'s (5//=, 5//@) study of 'causation' also adds to Wendt's (.99@) previous contributions to the scientific realist approach, highlighting shortcomings in the prevailing (empiricist) view of 'cause' in the discipline (e.g. Bicholson, .99=b). While those like Mervin focusing mainly on mainstream philosophy of science are criticised for their narrow arguments, scientific realists are

targeted for emphasising ontology too much (Brown, 1995; Hernoff, 1995, 1996b). Friedrich Ratzel, for example, insists in the primacy of epistemology, and sharp differentiation between the practice of metatheorising and the character of social theory affecting political practice (Ratzel, 1902a; Wright, 1990). Patricia & Haddeus Jackson attempts to step back from the specificities of these metatheoretical issues, providing rather an overview of how they affect research methodology (1996b). Jackson's study, together with an advanced textbook chapter by Uri and Wright (1990), are so far the most relevant attempts to treat the issue of metatheory at a more general level. They lack, nevertheless, any intensive effort to define 'metatheory' and its overall roles. They also do not map the defences and critiques of metatheoretical research in IR or weigh relative merits of the key claims on each side.

Another type of 'systematic study of theory' in IR refers to how theories fit into 'self-images' of the discipline (Offmann, 1990; Smith, 1990A, 1990B). We can find this connection between research on IR theory and self-images not only in analyses stressing solely cognitive aspects, but also in 'disciplinary history' (Duilhot, 1990a; 1990b). Schmidt, 1990a) or, alternatively, research on 'theory in context', including applications of the sociology of knowledge to IR (Taeber, 1990, 1991). Of particular relevance in recent publications is the combination of both history and sociology of IR theory into general frameworks (Lu'an D -ansen, 1999; Lu'an D Eittle, 1990). A central issue addressed in this vein is the 'Western-centric' character of IR scholarship (Inayatullah & Laney, 1990G; Eene, 1990). 'Eurocentric' limitations in the study of world politics are denounced and critiqued in a number of ways (Barvalho et al., 1990; Hobson, 1990; Ayaoglu, 1990), but studies evaluating 'peripheral' strategies of scholarship are increasingly relevant to the construction of a 'self-image' that shows some of the 'flip side of the coin' (Acharya, 1990; Aydinli & Attwells, 1990). One key feature of this type of study is its combination of metatheoretical issues with broader views of the discipline in terms of its institutional structure. This, however, is also its main shortcoming as a replacement for what this thesis purports to do. While we have much to gain from empirical analysis of these factors, 'theory' is only one of the core issues. A debate on the use of 'systematic study of theory' is omitted, since the focus on theory is subsumed to the broader disciplinary research interest (Alinia, 1990; Peterson, D & Tierney, 1990; Richard D Coyle, 1990). In this literature, therefore, we also cannot find any detailed general treatment of metatheoretical research and what it does for IR.

A final kind of metatheoretical literature worth mentioning here comprises journal debates on localised elements of metatheory in IR. Two discussions have recently been capturing the attention of key scholars in the field. The first has to do with the importance of the philosophy of science for our discipline. The inaugural issue of *International Theory* contains a piece by Buno (onteiro and <even Ruby (5//9a) launching an attack on metatheorising. It argues that applications of philosophy of science to IR have done much disservice to substantive research and cross-theoretical communication. Those who replied to the article raise a number of points also related to metatheorising. <urli (5//9) addresses the role of political ideology in shaping metatheoretical and theoretical positions, while &ac!son (5//9) defends the examination of philosophical assumptions dividing the discipline. -e partly agrees that ultimately these are not sufficient grounds for avoiding cross-theoretical dialogue. 6hernoff (5//9b) once more defends his conventionalist view and the exercise of employing philosophical tools in the discipline, but also analyses the internal logic of (onteiro and Ruby's argument. In their rejoinder, (onteiro and Ruby (5//9b) further clarify their position and defend a move to substantive and eclectic theorising as a result of what they call 'foundational prudence'. A second relevant debate draws on a similar perception that substantive research has been slowed down by metatheoretical divides (<at?enstein D il, 5//@). &aling this perception as a starting point, Rudra il and 3eter <at?enstein (5/./) attempt to show the added benefit of cross-theoretical 'syntheses' emerging from more openness at the metatheoretical level. In *International Studies Quarterly*, Cavid Ea!e (5/..) reiterates the point and adds his own view that a construction of a 'common le\$icon' to facilitate cross-theoretical communication could provide a platform for substantive research driven by a more pragmatic focus. -enry Bau (5/..) replies that, due to the nature of social science, metatheoretical 'drives' behind substantive theorising is unavoidable. ,or e\$ample, Ea!e's 'le\$icon' privileges mainstream theorising, rather than erasing the 'insulation' between mainstream and alternative approaches. il and <at?enstein (5/..), on the other hand, are relatively more sympathetic to Ea!e's critique of the disciplinary insulation of 'isms', but insist that their own version of 'analytic eclecticism' would be a more helpful approach to address the issue. +nce again, we see little commentary on the general function of metatheory in IR, but in-depth discussion on some of its aspects and consequences.

The relevant literature, then, often assumes a number of positions on the role of metatheoretical scholarship and sometimes vaguely defend these positions, but we still

lack a more general study on metatheory itself from the perspective of our discipline. It is mainly in response to these shortcomings that I claim originality for the present study. The assumptions adopted in this literature are, for most part of the time, implicit, although they do frame each author's approach in particular ways. This thesis, conversely, purports to unpack a number of points deriving from the notion of metatheory as 'systematic discourse about theory'. In order to do so, it (.) conceptually looks in detail at the nature of metatheory; (5) critically considers the IR literature both against and in favour of metatheoretical research in the field; (F) analytically elaborates an argument in response to these positions; (G) further addresses implications of metatheorising as a mechanism that shapes the discipline in specific cases. A detailed look at the structure of each chapter and how they connect to the main argument will help to clarify these claims.

The argument under consideration

The present study is, first of all, an attempt to provide a conceptualisation and classification of metatheory as it applies to IR. Chapter 1 sets the scene, so to speak, by mapping what 'theory' means to key scholars in the discipline, and by deriving a series of "questions which may only be addressed at the level of 'systematic discourse on theory'. In an intuitive way, we are thus confronted by a number of demands for metatheory. These intuitive demands are confirmed, next, in Chapter 5, where this basic definition of metatheoretical research is related to some of the few cases in which it figures in the IR literature. The same chapter expands this framework in light of similar studies in cognate disciplines (politics, economics, sociology) and provides a general typology of 'metatheory'. The term has also been differently employed sometimes in these correlate disciplines. For this reason, I further clarify the conceptualisation and typology with a short discussion of what I do *not* mean by metatheory. Specifically, I stress the ways in which metatheory differs from philosophy, although there is some overlap between them. The lack in IR of a general understanding of 'metatheory as such' (despite the sporadic appearance of short definitions and even detailed applications in specific issues) is only one research gap addressed in the first part of this thesis.

Another gap is the lack of an overview of what IR scholars have said about the roles metatheory has played in IR. In the second part of the thesis, the initial goal is to organise the literature in terms of the main claims against metatheoretical inquiry and the main claims in its favour. This is executed, respectively, in Chapters F and G.

Besides organising this material, I also provide a critical evaluation of both negative and positive views of metatheory. They often cancel each other out. (Moreover, within each view there are also opinions that seem inconsistent with each other. Finally, some of the points on either side may still be maintained. For this reason, in Chapter 6 I call for a "qualified defence of metatheoretical inquiry in IR. I do so, first, on account of the existing literature on why we need metatheory, how we should improve it and what we should avoid. Secondly, I provide my own 'analytical', deductive, argument to secure a place for metatheoretical discourse in IR. I prove that one cannot completely eliminate metatheoretical discourse from a discipline especially when one systematically argues against it. (My analytical claim is 'transcendental' in that I investigate what it takes to make a strong and absolute objection to metatheoretical discourse in this sense. I find a contradiction between these conditions of possibility and the actual argumentation for the absolute elimination of metatheory from a given discipline. Hence this full argument is ruled out, we are thus left with the proposition that there is, indeed, at least *some* room for metatheory in a discipline. We might as well look at its problems in the concrete situation of IR and try to contribute in such a way as to take into account the other, more plausible, objections raised by critics of metatheory in the discipline. What the second part of the thesis provides is a critical survey of the IR literature and a limited defence of metatheory in our field that allows for some criticism but ultimately acknowledges the relevance of metatheoretical scholarship.

We do well in bouncing the analytical and deductive argument back to concrete discussions in the discipline. The third part of the thesis, then, indicates how metatheory helps us analyse, assess and account for a selection of debates on world politics. In a series of 'cases', I address what metatheory is doing *for protagonists* in each debate, and what it can do for *us*, as we look at the cases. For each instance, I pick a central issue discussed in light of one aspect of the study of world politics, identify diversity of theoretical claims around that issue and identify roles played by metatheory in that case. I end each analysis with an examination of how studying *them* is illustrative of the general roles ascribed to metatheory in the initial part of the thesis. The first case, however, is somewhat different from the other two. Contained in Chapter A, it is an exercise in dialogue with International Political Theory (IPT), and addresses a 19th century debate on what the only Roman Empire was. The authors selected for this exercise (Johannes Althusius, Samuel Pufendorf and 2. 7. Leibniz?) have shaped modern political thought in different ways and are now recognised as relevant theorists

of sovereignty and federalism (Hulau, 1996). Althusius was selected not only because of the originality of his formulation, but also because of the historical context. Writing before the Peace of Westphalia (1648), which substantively altered the configuration of the Empire, he presents an interesting portrait of an already decentralised polity. Post-Westphalia thinkers Zuefendorf and Eiebnitz shared a similar background and metatheoretical assumptions and yet presented diverging accounts of the Empire. In combination, these differences and similarities leave us an interesting selection of examples. The chapter operates in this thesis as a 'control' case, whereby I establish the validity of claiming that metatheory shapes theory in this case, political theory in regardless of a well-defined disciplinary context. The first study, therefore, will mainly interest political theorists and historians of ideas.

Cespite its pre-disciplinary focus, the case can be of general interest also to IR scholars, considering that both the Empire and its constitutional transformation in the 17th century are at the centre of a key controversy in the history of International Relations related to how Westphalia shaped the early modern states' system. This latter theme is my second 'illustration case', in Chapter 2. I look at three distinct theoretical views of Westphalia in IR and their respective metatheoretical 'drives'. Their authors (Adam Watson, Daniel Philpott and Lenno Giesche) were considered of relevance in my design of this illustration due to not only the prominence of their treatments of this historical topic in IR, but mostly the divergence between the 'schools' of thought to which they belong. Watson's 'English school' approach is more inductive, teleological and comparative. Philpott's individualist 'soft constructivism' affirms the primacy of ideational factors and attempts to fit them into a 'chain of events' causal framework. Giesche's materialist analysis is influenced by Marxism and provides a critical, dialectic and constitutive account of systemic processes. I argue that these metatheoretical positions constrain and enable certain types of theoretical claims on 'order' and 'change', the nature of historical inquiry and how to execute it in IR. Metatheory also leads to specific dynamics across different theories on the same issue and sometimes can even account for lack of factual precision and help us understand why certain types of empirical data are selected in particular ways to support theoretical claims.

Chapters A and B provide in-depth analyses of a limited number of approaches to empirical and historical issues in one theorising the only Roman Empire and the other, Westphalia. Besides constraining and enabling claims in these debates, metatheory also shapes my own study. In both cases, it enables comparative interpretation of theory. In

Chapter A it also helps me formulate a historical, contextualised, account of political theory. In Chapter =, it provides the discursive space to engage in some evaluation as well. My main intention, however, is to illustrate the role it plays in the debates themselves, focusing more on the authors analysed in each chapter than on my own manoeuvres. For this reason, the 'self-reflexive' study of the overarching metatheoretical discourse analysing those debates is more implicit in those chapters than in Chapter >.

In dialogue with approaches in IR and International Political Economy (IPE), this final case looks at the notion of 'hierarchy' in the international system according to a number of 'research programmes'. Hierarchy is often mentioned in accounts of the transition allegedly marked by Westphalia in world politics. However, the concept is also empirically relevant given the post-Cold War intensification of the asymmetry between the United States (US) and remaining countries in our system. The portrait of world politics as hierarchical challenges one of the core theories in IR, Kenneth Waltz's neorealism. His structuralist and instrumentalist view of hierarchy contrasting with anarchy is the first notion analysed in this case. David Easton's mainstream model of 'dyadic hierarchy' has attained prominence in the discussion and is contrasted to Waltz's view. Less conventional are the choices for the remaining theories. Carlos Escobar's 'peripheral realism' places weak states at the centre of the analysis of hierarchical arrangements and possesses an inherently normative character. Hans-Jürgen Lauth's individualistic approach based on 'Austrian school' economics avoids dialogue with IR but challenges some of the received assumptions in IPE about the nexus between 'wealth' and rise to 'power'. In this final chapter I make deliberate reference to John Hobson's (1905) study of 'Eurocentrism' in the historical formation of international theory as a metatheoretical framework that helps me organise the material. In applying his study to a number of theories of hierarchy, I ask whether it can make sense of the inner dynamics in each of the approaches analysed. Moreover, I go further than Hobson and ask for each approach whether its respective theory of hierarchy is consistent with the metatheoretical claims defended by each author. The reflection on the overarching framework employed in my understanding of this specific debate (i.e. Hobson's study), therefore, is more explicit. On the other hand, the fact still remains that each approach is driven by peculiar sets of internal metatheoretical assumptions which are more specific than Hobson's typology of 'Eurocentrism'. Finally, it becomes clear that metatheory helps us not only by providing a reading of each approach to hierarchy in terms of their 'immanent' elements and 'outside' relation to Hobson's framework. In this study, it also

provides a way of critically dealing with the shortcomings of this framework! and perhaps complementing it with additional considerations.

The claims advanced in this final part of the thesis should be tested with reference to the 'empirical material', or texts, to which it refers. I try to provide an in-depth and clearly organised reading of these texts, and my views should be checked against that material. I sought to strike an even balance between known and lesser-known authors in my choice, to make the selection itself another original contribution. Despite this, ultimately the point is to limit the number of analysed texts to a reasonable amount that allows for the required depth. What I hope to achieve with the 'cases' in this third part of the thesis is a clearer understanding of metatheory as a mechanism that constrains and enables certain theoretical and disciplinary dynamics in IR. This notion that it does affect the discipline complements my analytical argument, developed in part two, that there may be problems with metatheoretical research in the field but we cannot completely eliminate it from IR. At that step of the argument, the test criteria have to do more with the internal coherence of my claims and to whether they do justice to the literature surveyed and critically examined. Part one, to which we now turn, begins to explore the general notion of metatheory as 'systematic discourse about theory' or 'scholarship about scholarship' by executing, first, a survey of many conceptions of 'theory' in our discipline. In IR, the level of debate at which the definition of theory, its forms, uses, evaluation and other issues operate is not the level of the study of world politics *per se*, but rather the study of *theory* of world politics. This 'second-order' character of metatheory is emphasised after the next chapter, where a typology of metatheoretical research and the roles played in general by each type is introduced and illustrated.

PART ONE

Chapter 1

Theory in the Discipline of International Relations

Being theoretical is unavoidable! "Why? Because the very process of engaging in observation requires sorting out some of the observed phenomena as important and dismissing the others as trivial! There is no alternative! The details of situations do not speak for themselves!"
*&Rosenau' () *' p!+, -!*

Introduction

(International theory has been provisionally defined as a systematic discourse on *theory*., and this leads us to the necessity of defining what is meant by 'theory' in IR. Writings of social scientists and philosophers of science are helpful here, not least because they have been influential in the discipline. &here are those who find this material wanting. Much writing on social scienceM, says mainstream IR theorist Stephen Han #vera (.99>, p.F), Lassumes that readers already know what theories are, what good theories are, what elements theories contain, how theories should be expressed, what fundamental rules should be followed when testing or applying theories, and so onM. Yet, practically every key IR theorist, economist or social theorist has made some kind of declaration on the nature and role of theory in the field. &hey may lack detail, as Han #vera implies, but we cannot deny their influence in shaping their respective disciplines. We can think!, for example, of Morgenthau's pamphlet differentiating theory in politics from theory in natural science, or Martin Wright's well-known essay on the distinction between international theory and 'domestic' political theory. &r maybe the debate between Hedley Bull and Morton Kaplan on whether IR theory should be 'traditional' or 'scientific'. We may think of Kenneth Waltz's stipulation that theories should be judged by their usefulness in prediction, and that they should be 'systemic', rather than 'reductionist'. &r maybe the widely cited textbook by King, Kohane and Herba demanding explanatory, causally-inferred theories from IR scholars. #Samples abound of such pronouncements *ex cathedra* from these central figures in the field. A similar pattern can be detected in other disciplines.

&o his credit, Han #vera is partially right. &here is an abundance of material on

theory in the social sciences, but most of the issues still revolve around what theories do, how to test them, etc. rather than what they *are*. In addition to this, when IR theorists *do* state in brief what they mean by 'theory', at the most 'basic' level the definitions tend to be "uite similar. &hey may not be worded in e\$actly the same way, but there is enough resemblance to allow for a 'common denominator' definition of 'theory'. imply put, a theory is an argument, a set of interrelated propositions about a sub9ect'matter. &heory is systematic discourse about something (see Iunge, .999, p.). 7hat usually happens, though, is that these authors "uic!ly move on to embed their respective definitions into a wider framewor!. &his is fre"uently done with reference to the philosophy of science or to social science, which might e\$plain the assumption that we !now the literature enough to 'fill in the gaps' with the remaining details. In the end, despite all the agreement on the 'basic definition' of theory, when it comes to its implications there is considerable diversity. #ven IR theorists belonging to the same school of thought might disagree on what the roles of theory are, how theory operates in the logic of scientific research and how we should assay theoretical formulations in the discipline. &a!ing that into account, I see! to delineate, whenever possible, their statements not only on the 'basic definition' of theory, but also on the roles and elements of theories, and how to evaluate them.

In IR, the wider intellectual conte\$t in which discussions of 'theory' emerge is often related to selfreflective 'debates' that have shaped the discipline (Ei9phart, .9>Ga; mith, .99A). &e\$tboo!s and review pieces tend to rely on the disciplinary narrative of the so'called '2reat Cebates' and a series of derived cross'theoretical dichotomies or 'disciplinary divides' (-olsti, .9@A). 7hile there are reasons to "uestion the framewor!, some familiarity with it is re"uired because IR theorists often refer to it, or assume it, when they portray 'theory'. &he ,irst Cebate allegedly divided IR between 'realists' and 'idealists'. &he narrative states that the !ey point of tension was whether world politics should be studied 'as it is' (realists) or 'as it ought to be' (idealists) (but see Ashworth, 5//5; chmidt, 5/.5). Idealists (particularly the liberals) were motivated by an 'ideal of control' over the international system and understood it in light of some sort of classical progressivist framewor! (2uilhot, 5/..b). Realists (particularly those drawing on political theory) were more sceptical of this ideal, given historical contingency and the comple\$ity of human nature and its 'will to power' (Ashworth, 5/..; earsheimer, 5//A). &his other aspect of the initial tension was made manifest in another 'divide' in the econd Cebate, between those defending IR as a 'hard science' and those adopting

practical reason, historical evaluation and ethical/legal judgement as the main guidelines for IR as an 'art' (Gurtis & Oivisto, 2005). The third debate, of lesser importance in this discussion on 'theory', still made reference to it, portraying theory in terms of 'paradigms' in competition (Lan!s, 1999A; but see Zaver, 1999=). It was widely believed that IR should eventually agree on a single 'paradigm' in order to move on with empirical research. The fourth debate "questioned once again the role of IR theory (Epid, 1999@). A new 'divide' emerged between those defending 'interpretive' theorising (the postpositivists) and those adhering to 'explanation' as its predominant role (often labelled as 'positivists') (Ollis & Smith, 1999=). Accounts of this wider intellectual context in which views of 'theory' have emerged in IR are deeply contested (Zaver, 2005). For one thing, it provides too uniform a view of 'postpositivism' that needs to be unpacked, for example, in the responses referring to 'critical theory', 'practical judgement' and 'incredulity toward meta-narratives'. Disciplinary historians have correctly dispelled the great debates story and the survey in this chapter will contribute to advance the cause of a more nuanced 'self-image' of IR (Zuilhot, 2005a). (My reference to this contested account is simply to point out its pervasiveness in the discipline.

This chapter outlines, first, some of the key positions in philosophy and social science influencing the debate on the nature and status of theory in IR. After that, I further detail the connection between this literature and some of the most important views of theory in IR. As it turns out, both mainstream and alternative approaches in the discipline owe much to philosophy and general social science. I deal first with classical, normative and mainstream IR and then with alternative views of theory in the field. This is followed by a brief discussion, in which I raise a number of "questions about the nature of *inquiry and theory* in IR, guiding us to the next chapter.

Context in philosophy and social science

Two philosophical ways of approaching science have influenced our discipline. The first looks at isolated theories and their empirical subject-matter. The second approach looks at theories in a more dynamic way, as part of the 'growth of knowledge'. Discussions on 'syntactic' and 'semantic' views of theory, logical positivism and falsificationism refer to the first approach. The 'syntactic view' of theories (Bagel, 1999=; Zartofsky, 1999@) sees

Some will number this debate on postpositivism as the third debate, but I follow Zaver's (1999=) account of it as the 'fourth debate' in IR is also a popular narrative.

them as collections of statements that can have a formal representation as axiomatic systems. The logical structure of a theory is separable from its factual content (Iortolotti, 5//@, p.AA). It would, therefore, be possible to distinguish between 'uninterpreted' implications of axioms from 'interpreted' logical, observational and theoretical terms. This notion was very popular in the golden age of logical positivism, and widely subscribed by scientists.⁵ Logical positivists had a very restrictive view of theories as part of 'the scientific method', in which only deduction and empirical verification of hypotheses relying on the existence of 'brute facts' would count as knowledge (but see Hempel, 9=A). Having emerged in opposition to the oversimplifications of this view and its almost exclusive concern with formal aspects (Achinstein, 9=G, p.F5@), the 'semantic' account focuses on the role played by theories as models of reality (Uppé, 9>G). The practice of representation is key to our understanding of science (Ziere, 5//G, p.>GF). In fact, some proponents of this alternative approach go as far as to reduce theories to models (e.g. Ziere, 9@@, p.@=), or at least deny the primacy of a formal language (van Raassen, 9@/, pp.GG; =>), without being altogether hostile to a certain degree of formalisation (Lunge, 99@, p.A/F).^F For others, there is a qualitative difference between models and theories, the former being employed to formally represent not only the subject-matter but also the theories themselves (Morrison, 5//>, pp..9A'=).^G While 'syntacticists' highlight the inner structure of scientific theories, the 'semanticists' focus on the relation between theoretical models and the empirical subject-matter.

Classificationists, following Carl Popper, also emphasise this relation, albeit in a different manner. Popper and influential positivists subscribe to the 'deductive-nomological' and the 'hypothetico-deductive' descriptions of scientific research. The deductive-nomological model forms theories by extracting implications from general premises where at least one part of the starting point is a 'law-statement', i.e., a

5 Albert Einstein, for example, defined theories as 'the totality of the primary concepts, i.e., concepts directly connected with sense experiences, and theorems connecting them'. Of paramount importance in his view is 'the aim to represent the multitude of concepts and theorems, close to experience, as theorems, logically deduced and belonging to a basis, as narrow as possible of fundamental concepts and fundamental relations which themselves can be chosen freely'. It is 'by means of such concepts and mental relations' that we may 'orient ourselves in the labyrinth of sense impressions' (5//=, p.AF; A=>).

F Contrary to Rigg and Hartmann's (5//9, p.5F) opinion that the semantic approach requires the rejection of formalisation, advocates of this notion have defended even axiomatic formalism under certain conditions (Achinstein, 9=@, pp..AF'G).

G Dichotomising between 'syntactic' and 'semantic' notions of theory in a very strict sense is unhelpful. There is more potential dialogue across the approaches than their respective advocates would admit (Endry D Sillos 5//>) and, besides, as we shall see below, in the case of IR, scholars do indeed acknowledge theories as being both 'arguments' and 'models'.

description of a pattern of regularities and correlations. It is also called the 'covering law' model, because predictions are made based on what can be inferred from the law' statement, which is supposed to 'cover' or subsume singular occurrences as an instance of the law stated (Sillars, 1978, pp. 10-11). The hypothetico-deductive model, in turn, starts from the formulation of a 'conjecture' or hypothesis with implications and then empirically tests it (Sillars, 1978, pp. 11-12). Inductivist positivists, argues Popper, face the problem of demanding too much of 'verification' in a theory test. They cannot get around the fact that confirming the same hypothesis over and over again will never strictly *prove* it.^A Starting from 'if P then Q ' and finding N does not necessarily mean that P obtains. Instead, conjectures should be tested with a view to proving them *wrong*, i.e., 'falsifying' them% 'if P then Q . But Q , therefore, not P ' (Popper, 1959A). Notice that both approaches assume a very narrow idea of 'causation' 0 they reduce 'causation' to a discrete relation between P (cause) and Q (effect), as if everything operated like colliding billiard balls. In this 'chain of events' account, causes are merely sequences of correlated events (see Achinstein, 1965). For falsificationists, science can be demarcated from other forms of discourse by being open to criticism (Popper, 1959B). Attempts to 'immunise' a falsified conjecture either by adding *ad hoc* hypotheses covering some of the challenging evidence or making the theory unfalsifiable altogether are pseudo-scientific.

Parallel to this discussion is the issue of 'instrumentalist' and 'realist' understandings of theories. Instrumentalists treat theories primarily as tools, 'maps' or 'filing systems', designed to organise and connect statements about observation (e.g. Cuhem, 1968). Theoretical statements should be evaluated not in terms of their 'realism'. Rather, we should ask whether they are 'useful' in this 'organising' and 'simplifying' role, as well as in their capacity to predict. H. P. Nune challenged the empiricist stance defended by positivists, including their view of 'brute facts', by arguing that hypotheses cannot be tested in isolation, given the possibility of providing a number of equally reasonable 'maps' of a given 'pool of facts'. Some had previously argued that hypotheses should necessarily be placed in a wider system, a theory (see Jelic, 1969). Admitting the 'theory-ladenness' of scientific observation (2odfrey's myth, 1976, pp. 10-11), Nune (1969) went even further and claimed that a clear-cut

A Inductivism is the doctrine that stipulates that knowledge of the general may be obtained from the examination of a great number of particular instances. The 'problem of induction', described here, had been formulated long before Popper, in the epistemological writings of empiricist philosopher David Hume.

distinction between 'propositions' and 'facts' against which they should be 'tested' is ultimately untenable. It is the field of science that should be tested as a whole, and not isolated theories. For instrumentalists, therefore, the general purpose of science is to make useful predictions and give us manipulative power. Theories, in their condition as instruments, can be assayed in terms of their utility, not 'realism' (Cartwright, 1983, p. 105). Realists, on the other hand, regard theories as attempts to describe reality even beyond the realm of observable things and regularities (Birnbaum, 1969). This does not necessarily mean that realists discard the 'mapping' role played by theories (Cartwright, 1983, p. 105). Science, says a leading scientific realist, tends to build conceptual mappings of the patterns of facts, i.e., factual theories. But theories are more than maps. They establish logical relations, provide explanation via interconnected hypotheses and enhance the testability of hypotheses taken together (Lange, 1990, pp. 105-106). Explanation, in the realist sense (and deeper than the positivist approach of 'covering law' in light of 'correlations'), means accounting for underlying structures and mechanisms which are involved in causal processes (Cartwright, 1983, p. 105). Realists argue that a key role of theory is to uncover these mechanisms, and that successful theories will contain models that bear some correspondence to reality (Lange, 1990, p. 106). It is important to acknowledge this parallel debate because, first, it goes to the core of the issue of theories and their role in science. Secondly, some of its aspects like the 'theory-laden' character of observation also recur in general social science and IR. As we shall see, a collage of positivism, falsificationism and instrumentalism has grounded some of the mainstream positions on theory in IR. Realist views have also emerged in the field.

These three engagements – syntactic or semantic, positivism or falsificationism and instrumentalism or realism – have something in common. They tend to address isolated theories in relation to their empirical subject-matter. Following Nuñez's thesis, alternative accounts of theory in the philosophy of science, in turn, understand them as part of the 'growth of knowledge'. That is to say, embedded in wider frameworks. Notions like 'research programmes' and 'paradigms' are part of this second approach. We shall look into these in more detail when we address their respective IR adaptations. (Meanwhile, suffice it to say that their original elaboration, aiming at an understanding of natural science, has been extrapolated and transposed to IR. The same is the case for the philosophical literature on theories in isolation. Philosophy and, mainly the philosophy of science, is the first type of source scrutinised by IR in its quest for a more

clearly articulated notion of theory.

Another intellectual milieu that has been considerably influential in our discipline is that of general social science. (ainstream IR, under the influence of both natural science and economic methodology, relies more upon an essentially 'unitary' view of science as a whole. ome of the alternative views of theory in IR and social science highlight the inherent split between natural and social worlds, where ideas constitute to a great e\$tent the sub9ect'matter. &herefore, we should be open to a number of reconfigurations of IR as an 'interpretive' study of society (rather than part of a 'unified science'). ,irst, we should be aware of the impossibility of value'neutrality, as values are both part of the worldview of the researcher and part of the sub9ect'matter.

econdly, we need to tale the critical potential of theorising more seriously. &hat is, we need to loo! at the social world not as given, but as constructed. 6onfigurations of power and in9ustice, therefore, are not fi\$ed, but malleable to change. &hirdly, 'refle\$ivity' means that scholarship about the social world also potentially helps to *shape* it. 7e cannot postulate a sharp distinction between researcher and sub9ect'matter. Coing social science means constructing the social world. Among the non'mainstream proponents of IR as a social science, some still defend an e\$planatory role of theory, but in terms of 'constitution' and 'mechanisms', while others prefer to ma9or on 'refle\$ivity' and 'interpretation' and see an unbridgeable gap between 'e\$planatory' and 'interpretive' roles of theory. &here are also those who re9ect the pro9ect of 'IR theory' in a scientific (even social'scientific) fashion. +thers tale it a step further and denounce the e\$clusionary drive of theoretical discourse as power capable of imposing 'order' upon the world. In what follows, each position is discussed 0 IR theory as social theory and IR theory as discourse and social practice. &hese two, however, follow a study of classical, normative and mainstream views of theory in IR.

Theory in classical! "eha#ioural and normati#e IR

-ere I 9u\$tapose classical and normative IR because of the clear continuity between their views of theory's definition, elements, use and assessment. Additionally, there is also continuity in that both articulated their position on the conduct of in"uiry against claims to 'science'. 6lassical or 'traditionalist' IR is said to have pioneered the field alongside 'scientific' technocratic'liberal attempts to e\$plain and control public policy at the international level (2uilhot, 5/..c; Nuir! D Higneswaran, 5//A). #ventually, it formulated a clear reaction to the rise of behaviouralism in political science and IR.

Normative IR, later on, deliberately drew on traditionalism and occupied its disciplinary space in reaction to mainstream IR. I tackle mainstream IR in the next section.

Theory and ethical judgement was defended in the classical fashion against particular forms of 'raw empiricism' advocated by some of the new 'scientific' approaches under the influence of inductivist positivism.⁵ In behaviouralist IR, world politics was undertheorised and allegedly 'value-free'. Instead, isolated (and often contradictory) hypotheses were elaborated and 'tested' against empirical evidence in a 'neutral' way, seen as devoid of assumptions (see Hincant, 1976, p.10). Due to such efforts, databases like the *International Project* (Singer, 1975a) were created, providing a catalogue for data. Other behaviourists in politics defended some role for theory, although it was very instrumental and limited (Kaplan, 1977). For example, argued that theory is just a tool to transform facts into statements that should be tested in empirical research. The overall study of politics was portrayed as something built by the slow, modest and piecemeal cumulation of theory, methods and data (Hulst, 1976, pp.10; ..). The 'neutrality', in any case, was a key principle advanced by behaviouralist IR.⁶ Ethical judgement was replaced with 'policy prescription' in that framework. And be sure, scholars like Hulst (1976) himself could still write *political* theory in a more traditional fashion, but *empirical* theory should be kept separate from these philosophical endeavours. As a consequence of the 'behaviouralist revolution' in the general study of politics, then, 'scientific' IR developed an idiosyncratic framework for research (Haston, 1979; Zunnell, 1976, p.95). It was stressing the centrality of theory and ethical judgement that traditionalists reacted.

Scholars in the classical tradition, such as Morgenthau and Ledley Hall, emphasised several interrelated types of political discourse beside common sense, practical, theoretical and philosophical knowledges. They markedly adopted a 'humanistic' and 'vocational' approach to IR in opposition to the 'scientific' and 'technocratic' view (Haston, 1976, pp.10/11). Morgenthau, for example, argued that IR should be studied as an art (1976, p.10)

Politics must be understood through reason, yet it is not in reason that it finds its model. The principles of scientific reason are always simple, consistent and abstract; the social world is always complicated, incongruous, and concrete. Politics is an art and not a science, and what is required for its mastery is not the rationality of the engineer but the wisdom and the moral strength of the statesman.

⁵ See Hincant (1976) for a discussion in the context of general political science.

⁶ But should we say, a key *value*?

Due to the complex character of politics and the decision to study it in this 'humanistic' fashion, the classical way of theorising was closely connected, as Hall (1999, p. 10) explained, to the study of philosophy, history and law. It relied upon the exercise of judgment and went beyond the strict standards of verification and proof required by the 'scientific' view, humbly recognising the role of intuition and the tentative and inconclusive status of the exercise. Reinhold Niebuhr (1952, p. 55), another classical theorist, defended a similar view based on the notion that the historical realm is not comparable to the realm of nature. IR theory, in his view, should be aware of historical fate. Morgenthau (1966), conversely, assumed continuity between both realms, but still defended classical theorising on the grounds that social reality is incredibly more complex and demands ethical judgment. Despite its critique of the rampant empiricism in the 'scientific' behavioural approach, classical IR did not dissociate theory from empirical evidence. Williams & Ruggie (1967, pp. 15-17) noted that empirical information is important, but we can only use it in light of some theory, implicit or explicit. In other words, a theory is a tool for understanding and brings order and meaning to a mass of phenomena, treating data in propositions of both objective and general validity (Morgenthau, 1966, pp. 66-67; GA5). IR scholarship is, therefore, inherently theoretical even if understood in this 'humanistic' fashion. Only later, to mainstream IR's merit, would a clearer notion of the crucial role of theory in research be articulated and absorbed by those interested in practising IR as a science.

As an outcome of the clash between classical and scientific scholarship in the discipline, in some settings IR has been studied in the line of political or 'normative' theory. Normative IR recovers the 'traditionalist' approach and takes issue with the mainstream on two accounts. On the one hand, it challenges the notion of value neutrality, which 'scientific' IR seeks to preserve. On the other, most of the mainstream theories assume an 'ethical vacuum' at the international level. Normative theorists side with classical IR on both issues. Searching for great works of international theory, Martin Wright famously defined it as a tradition of speculation (1999, p. 10). He denounced the lack of ethical imagination of the field, which tended to assume an inner connection between morality and sovereignty. Because of the lack of a supra-national sovereign, most of the classics of 'domestic' political theory could not be 'transposed' to the international level, hence the absence of an international 'mirror' of political theory (but see Suganami, 1999; Waller, 1999, 5/10). Now that the dichotomy between

KinsideP and KoutsideP of the sovereign state is deeply contested in the field (Zaller, 1999), there seems to be more room for philosophical and normative argument concerned with issues of justice, freedom, equality and how human beings can achieve the good life (Schmidt, 1995, pp. 9; 15).

Those who pursue this mode of inquiry in IR will readily admit that there are many ways of theorising world politics, even within the normative field (Gochran, 1999, p. 5; Rengger, 1999, pp. 10-11). Some, like Chris Brown (1995, p. 6), go as far as to affirm that a non-normative empirical theory is possible in principle, although rarely seen in practice, despite mainstream claims to value-neutrality. (Most, however, share with (Ervin, 1999, pp. 1-2; FG) the perception that there is no way in which social scientists may legitimately avoid becoming involved in normative theory. All theory of International Relations says (Collyer, 1999, p. 1). This normative theory, although only a relatively small number of works display awareness of the fact (see Rengger, 1999b, p. 10). As we shall see later, some normative theorists like Andrew Einstein (e.g. 1999) combine ethical thinking with 'critical theory'. In doing so, they are also interested in denouncing injustice in the current order and theorising alternative orders. IR theory defined in this ethical sense, therefore, does not necessarily exclude alternative ways of theorising, although a clear argument about the role played by values and ethical judgement is desirable.

As implied by Kohn (1999, pp. 10-11; AA), international political theory aims at exploring moral expectations, decisions, and dilemmas in world politics. Going beyond mere prescription, political theorists are mainly concerned with moral norms, frequently in order to revise and transform them. A similar definition adds specific direction to normative IR theory, which takes as its subject matter the criteria of ethical judgement in world politics and seeks shared principles for extended moral inclusion and social reconstruction in international practice (Gochran, 1999, p. 5). An example is in order. One of the most widely cited works in normative IR is Michael Zaller's *Just and Unjust Wars*. In that book, Zaller (1999, p. 10) looks at ethical claims about war, given the present structure of the moral world. His aim is not only to account for the ways in which men and women who are not lawyers but simply citizens (and sometimes soldiers) argue about war, but also to expound the terms we commonly use.

But that's not to suggest that we can do nothing more than describe the judgments and justifications that people commonly put forward. We can analyse these moral claims, see out their coherence, lay bare the principles that they exemplify. We can reveal commitments that go deeper than partisan allegiance and the urgencies of battle (O). And then we can expose the hypocrisy of soldiers and statesmen who publicly acknowledge these commitments while seeing in fact only their own advantage (O). We are rarely called upon to invent new ethical principles; if we did that, our criticism would not be comprehensible to the people whose behavior we wanted to condemn. Rather, we hold such people to their own principles, though we may draw these out and arrange them in ways they had not thought of before (Walzer, 1993, p. 11).

In this passage, Walzer details the elements and aims of normative theorising. A similar notion is advanced by Brown (1995, pp. 10-11), who tentatively defines political theory as 'the study of the search for justice in society' and applies it to IR in terms of a 'body of work' which addresses the moral dimension of international relations and the wide 'questions of meaning and interpretation generated by the discipline'.

In Brown's definition, the transformational agenda drops out and a hermeneutical concern is highlighted. In addition to this, there is an implicit historical feature. While not strictly necessary, the interaction with canonical texts on similar normative issues is illustrative of international political theory (Thompson, 1996; (Walzer, 1993), hence the relevance of the interpretive element (see Brown, Bardin, D Rengger, 1995, p. 11). Because of these features, normative theorists defend some criteria of evaluation for their research that would not necessarily apply to empirical theorising. Those arguing for a particular interpretation of a certain author or tradition of thought may refer to hermeneutics (e.g. Schleiermacher, 1996) or specialised works in political theory dealing with the issue of textual exegesis (e.g. Gully, 1996). Those doing normative IR theory in the broader sense may struggle to find uniform criteria, although clarity and logical validity are generally seen as positive properties of good normative thinking (Rawls, 1993). An ability to deal with specific cases in the present or the past is also desirable (Walzer, 1993, p. 11). Traditional approaches to international ethics may have sought to provide a wide foundation in general ethics in order to ground specific claims on world politics (e.g. Bennett, 1996). (Most of the contemporary views, however, tend to side with Walzer (1993, p. 11) on the specific applied focus.)

I am not going to expound morality from the ground up. Where I to begin with the foundations, I would probably never get beyond them; in any case, I am by no means sure what the foundations are (O). The study of judgments and justifications in the real world moves us closer, perhaps, to the most profound 'questions of moral philosophy, but it does not require a direct engagement with those 'questions.

Like key to this theoretical practice, for several normative IR theorists, is finding a method by which one can discriminate between competing conceptions of international justice without falling foul of the charges of political and philosophical idealisation (Mutch, 1997, p.59). This is a criterion particularly designed to evaluate normative theorising.

Science and theory in mainstream IR

The mainstream reflects a sharp, qualitative, distinction between social and natural science. Differences are, rather, a matter of degree (e.g. in objectivity, accuracy, measurability, and so on). Much mainstream social science writing on the role of theories in the conduct research has adopted a combination of logical positivism, falsificationism and instrumentalism (Lyant, 1994, pp.11-12). Milton Friedman's influential essay on economic methodology is a case in hand. He trusts 'positive' economics (as opposed from 'normative') can be studied in an objective, value-free fashion in precisely the same sense as any of the physical sciences (1953, p.6). For him, the ultimate goal of a positive science is the formulation of theory, comprising two aspects. On the one hand, it is a set of tautologies that organises empirical material. On the other, it is a body of substantive hypotheses designed to abstract essential features of complex reality. We judge the former via formal logic, whereas the latter should be evaluated by its predictive power as an explanation for a class of phenomena (1953, pp.13-14). Up to this point, his argument follows the logical positivist view of theory. Predictive power, however, is understood in falsificationist terms: if prediction fails, the conjecture needs to be adjusted. Despite this, it is in an instrumentalist fashion that Friedman breaks down the process of evaluating theories. He asserts that, in terms of their assumptions, theories must be judged not by their 'realism', but rather by their 'usefulness' in prediction (1953, pp.13-14). Moreover, theories are constructed in a creative act of inspiration, intuition, invention (1953, pp.14-15). Friedman's case is illustrative of the blending that occurs when the philosophical literature is adapted to social science.

Mainstream philosophy of social science handbooks have indicated a widespread adoption of this type of hybrid approach. Like structural characteristics of a social science theory, says Richard Rudner (1965, p.1), are precisely the same as those of any other scientific theory (O). A theory is a systematically related set of statements,

@ At times, Friedman is somewhat vague and conflates 'theory' with 'hypothesis' and 'model'.

including some lawlike generalisations, that is empirically testable. While formalisation is something social scientists should aim for in their theorising, this is not always possible or even desirable. Anyway, *fruitfulness* and *simplicity* are criteria we should employ to judge social research (1999, p.6). Gerald Ager (1995, p.15) comments on a general agreement in sociology that a theory is a set of propositions or theoretical statements, consisting of both *concepts* and their connections. In terms of evaluation, the suggested standards echo Friedman's instrumentalism: *scope*, *parsimony*, *precision of prediction*, and *accuracy of explanation* (1995, p.15).⁹ Later authors indicate the adoption of a more nuanced view in mainstream social theorising. William Lidmore, for one, opts for a 'basic' definition of theory as *systems of thought* which in their presentation could be arranged in a *continuum* between formal and informal (1994, pp.15). Among the uses of theory are the organisation of a *chaotic mass of data*, establishing *meanings* for concepts and constructs, and *relating* somehow a conceptual problem or set of observations to a theoretical construction of reality which fits it. The latter is *explanation*, which is *the main goal of sociological theory* (1994, pp.6-7). The author further notes that bias should be avoided, and we should theorise about *things* which in principle can be measured, counted, observed, and correlated, although he acknowledges the inevitable presence of some degree of *subjectivity* and *ineffectiveness* (1994, p.5). Whether in a nuanced manner that accepts other ways of theorising, or in a more dogmatic way, mainstream social science has conveyed not a 'pure' positivist, falsificationist or instrumentalist view, but rather a *mixture* of these elements. This inclination to hybridise, in turn, was absorbed by IR and is manifest in the works of scholars like (Michael Bicholson, Robert Cochrane, Stephen Hanover and Kenneth Waltz).¹⁰

9 Mainstream views of theoretical enterprise in social science have brought discontentment to researchers due, perhaps, to the 'static' character of this view of theories, taken in isolation against empirical evidence (2005, 1999). Alternative interpretations of theories have been sought in the mainstream, still based on the philosophy of science. In economics, for example, these additional views are often set forth with reference to Popper's falsificationism (Haug, 199F), Eakins' methodology of research programmes (see Hands, 199F), Kuhn's paradigmaticism (Goats, 1999) and, more recently, scientific and critical realism (Eawson, 1999; (H1, 199@). There is something striking about most of these 'adaptations' of the philosophy of science to economics. Just like Friedman's instrumentalism, they tend to add elements of what economists perceive to be peculiarities of theory building and theory appraisal in their discipline, as well as traditional narratives in the history of economic thought (Lachouse, 199G; Caldwell, 199G; Eatsis, 1999). Whether these adaptations do justice to the original material in the philosophy of science is beyond the scope of this study, but the point remains that resulting frameworks on how to analyse economic theorising tend to be hybrids, mixing disciplinary peculiarities with the philosophical material (e.g. Haug, 199A).

10 It is telling that, due to such a strong preference for a 'unified' view of science, mainstream theorists prefer to hybridise this material rather than look elsewhere.

Bicholson's (.995, p.FG'F9) position is perhaps the most well-known version of what has been labelled 'mainstream positivism' in British IR (see also .99=a). -e describes theories as 'maps' to aid research. ,acts are only facts if seen in light of a conceptual framewor! that connects and organises them. trictly spea!ing, theory is only one part of the 'hypothetico'deductive' method of natural and social science whereby hypotheses are deduced from an initial set of a\$ioms and then sub9ected to empirical tests (.995, p.F/). *nli!e isolated hypotheses, theories have a general character% they are collections of Lpropositions which apply to classes of events, and not 9ust to individual eventsM (.995, p.5=). Interestingly, the author has labelled his position Lbroad positivismM and uses the term LreluctantlyM, re9ecting the Lbehavioural traditionM normally attached to it (Bicholson, .99=a). <ing, <eoane and Herba (.99G, pp.F' 9) follow Bicholson in the assumption of a single logic of empirical research for both natural and social science. LA social science theoryM, they say, Lis a reasoned and precise speculation about the answer to a research "uestion, including a statement about why the proposed answer is correct. &heories usually imply several more specific descriptive or causal hypothesesM (<ing et al., .99G, p..9). ocial science not only aims at describing, but also at e\$plaining by means of causal inference, generalising from a given collection of correlations. Ei!e Bicholson, they postulate that the focus of theories lies on generals, not particulars, accepting the validity of the hypothetico'deductive model. A theory is Ldesigned to show the causes of a phenomenon or set of phenomenaM, always containing an Linterrelated set of causal hypothesesM relating variables with Lobservable implicationsM to be tested (.99G, pp.FG'FA; 99'./). Han #vera (.99>, pp.>'9) provides a similar definition, but adds his own detailed emphasis on e\$planation. L&heoriesM, he says, Lare general statements that describe and e\$plain the causes or effects of classes of phenomena. &hey are composed of causal laws or hypotheses, e\$planations, and antecedent conditionsM. ly 'e\$planation' he means statements ma!ing reference to a chain of cause'and'effect events, observed as 'variables'. 7alt?'s (.9>9, pp.5'A) position echoes some of the above, although he stresses the role of creative conceptualisation in theory more than other mainstream writers. &heories are not merely Lcollections or sets of lawsM but also Lstatements that e\$plain themM. cience, whether natural or social, goes beyond description, creatively inventing 'theoretical notions' to circumscribe a field, and adopting them as a framewor! to organise data, e\$plain and predict (7alt?, .99/, pp.55'5=). 7hile sharing these aims of e\$planation and prediction with mainstream colleagues, 7alt? (.9>9, p...) points out

the need to go beyond the necessarily barren hypothetico-deductive approach^M, in that we need to understand theories as more than just the usual 'positivist' combination of inductive and deductive procedures, also including 'theoretical notions' and creative insight.

Waltz does not conclude his argument at the mention of 'theoretical notions' which are not 'falsifiable' in the Popperian sense. ¹⁰ Until this point, he is echoing positivism, instrumentalism and Friedman's hybrid economic methodology.¹¹ Later on, he takes a step further. In search for a more dynamic approach capable of dealing with the 'growth of knowledge' and theory appraisal, he connects his view of theory to Imre Lakatos' concept of 'research programme' (Waltz, 1979, pp. 17-18). According to Lakatos (1976), each programme contains an unfalsifiable 'hard core' of basic propositions (Waltz's 'theoretical notions'), defining its essential assumptions. It also contains a negative and a positive 'heuristic' and a 'protective belt' of 'auxiliary hypotheses'. The negative heuristic forbids certain amendments to the research programme, while the positive heuristic stipulates how one should proceed, including which ways of theorising would be compatible with that programme. While a programme is a large-sized and large-scoped theoretical system (e.g. Waltz's neorealist theory of international stability), it also contains additional theories of a more modest character (e.g. Waltz's account of emulation and socialisation generating life units in the international system), designed to protect it. In hindsight, we could 'rationally reconstruct' an episode in which one or more programmes developed new theories in response to empirical challenges (Qahar, 1976). A programme dealing solely with the challenging evidence, or going against its own heuristics, would then be labelled 'degenerating' and, in the long run, this would suffice to account for its loss of adhesion. A rival programme able to cope with the challenge and also predict 'novel' facts, without disrespecting the boundaries set by its heuristics, would be, over time, a 'progressive' and promising programme (Lakatos, 1976). By hybridising his view of theories with the 'methodology of scientific research programmes', Waltz was able to go further than his colleagues in the mainstream in terms of providing a more dynamic view of the 'growth' of theories and their appraisal.

After Waltz, a number of mainstream IR theorists have employed the methodology of scientific research programmes in order to portray their own theory as

¹⁰ Waltz references Ernst Bagel's 'syntactic' account of theories, but there is no citation of Friedman's essay on positive economics. Eater writings indicate his indebtedness to economic science (see Chapter 2 for more).

'progressive' and disqualify rival approaches as 'degenerating' (Herman D Herman, 5//5, 5//F). As in Walt's case, most often these transpositions from the philosophy of science to IR have resulted in modifications of the original material in order to accommodate it to the discipline (E. 2. Eire, 5//@). Another 'dynamic' formulation of interest to mainstream IR is Thomas Kuhn's view of 'paradigms' and 'normal science'. Kuhn (1996) understands theories in terms of two main cycles in the historical formation of a given scientific discipline. 'Normal science' predicates research on the dominance of a given 'paradigm', or a set of directives, implicit or not, on which scholars operate. Although we may find it hard to pinpoint an accurate definition of 'paradigm', it is often seen as revolving around a hegemonic theory, also recommending methods, determining relevant questions and ways of doing research.⁵ The other cycle of a discipline is connected to 'revolutionary science'. Contrary to what Popper says, if a paradigm is strong enough, it can afford to ignore falsifying pieces of evidence. If it persists ignoring them, though, scientists may feel that the approach is losing its explanatory power and may resort to an eventually available alternative theory (Dodgson Smith, 5//F, pp. 46ff). Mass adherence to a new paradigm leads to a qualitative shift, a 'scientific revolution'. From a Kuhnian perspective, the point of science would be to find an agreement on general paradigmatic principles as soon as possible and then move on solving puzzles on that basis.

selective readings of Kuhn prevail in IR. Mainstream theorists prefer this version to a normative claim of 'normal science' being the aim of IR. If we agree on a uniform 'paradigm', we will eventually go to the next stage and solve puzzles (Einhart, 1997b). Others, aware of the relativistic implications (addressed later in this chapter), prefer to blend Kuhn's 'paradigmatism' with additional material – usually a mixture of Popper and Ekeland (Hastings, 1999A, 1999). This sort of move has led to some controversy, not least because reference to Popper, Ekeland and Kuhn tends to imply assumptions about how research should be conducted in IR, and what makes a good theory. For instance, in a journal controversy that attracted some attention, John Hastings was questioned along these lines by Walt (1997), Randall Schweller (1997), Stephen Walt (1997) and other colleagues (Christensen and Snyder, 1997; Herman D Herman, 1997). Hastings (1997) had to defend not only his 'hybrid' framework and its implications for theory appraisal, but also debate whether his reading of Ekeland was adequate, and whether his 'rational reconstruction' of Walt's neorealism as a 'research

⁵ On the ambiguity of the term 'paradigm' in Kuhn's work, see Asterman's (1997) discussion.

programme' was appropriate. For the latter scholar, curiously, the methodology of scientific research programmes is a way of drawing a line between his own approach and empiricist positivism (Waltz interviewed by Allan D Rosenberg, 1999).

So far, we have looked at classical, normative and mainstream views of theory, its definition, roles and assessment in IR. Both classical and normative IR draw on general political philosophy and correlate areas in order to argue for the indivisibility of ethical judgement from IR theory. They both postulate a set of peculiar conditions separating political practice from the subject-matter of the natural sciences. They also highlight the need for theory, an element neglected to a large extent by the 'scientific' proposal of behaviouralist IR under the influence of positivism. Drawing on additional elements in the philosophy of science, social science and economics, mainstream IR embraces a more nuanced view of theories. Some scholars look even further at the philosophy of science in order to incorporate a more 'dynamic' view of theories embedded in 'research programmes' or 'paradigms'. In all these cases, however, a new formulation emerges from the decontextualisation of the original approaches. Still, all of them assume some sort of continuity between natural and social science that would enable the move. Those who understand otherwise look at other sources to support their view of IR theory as social science or, alternatively, as social discourse or practice.

IR theory as social theory

Mainstream views of theory in IR delineate its role as primarily one of explanation as a 'covering law' and 'causal inference'. A certain event is explained as an instance of a statistically grasped 'general law'. Its 'causes' are identified in correlations and 'chains of events'. This view of the role of theory has been challenged in social science in a number of ways. We turn, now, to views of theory in IR predominantly influenced by alternative approaches to social science, based on 'interpretive', 'critical' and (scientific) 'realist' perspectives. 'Anti-naturalists' postulate an intrinsic separation between the 'natural' and 'social' realms. The manner in which social science reflects this distinction is in the aim of employing theory to 'interpret' human action. Interpretivists in this sense either reject the notion of 'explanation' altogether or attempt to reform it in light of alternative insights. There are also those who reject the mainstream view but still adhere to a 'naturalist' position. They tend to keep explanation as a key role of theory. Scientific realists and/or some of the critical theorists (in a broad sense) are the most common example in IR. Instead of adhering to the mainstream view of explanation, or rejecting it

altogether (as some 'interpretivists' do), they redefine explanation. A notion employed in this redefinition of the goal of social theorising is that of accounting for 'constitutive' causation, or perhaps for 'mechanisms'. This alternative view also makes room for 'interpretation'. It could also incorporate relevant insights on 'reflexivity' and the critical character of theory.

Let us look, first, at the interpretive views. Interpretive social scientists have adopted an anti-naturalist framework of 'social science' based on the idea of an intrinsic separation between the 'natural' and 'social' realms. By implication, theory in social science necessarily operates in a different way. Charles Taylor (1994A, p.95), for instance, laments the constant temptation to take natural science theory as a model for social theory and notes that the latter is part of a significantly different activity. (As Weber and Schmitt actualised this 'anti-naturalist' stance via an 'interpretive' or 'hermeneutical' approach to social theory. Weber (1985, p.55) in principle did not exclude 'explanation' from the goals of social research, but affirmed the need to go beyond an account of the 'chain of events' leading to a certain effect. Rather, an interpretation or *understanding* of the 'reasons' for social action in its unique setting should be provided. Understanding in this sense, said Weber, consists in placing the act in an intelligible and more inclusive context of meaning (1985, p.55F). The theme reappears in Schmitt's (1994A) influential work. According to him, doing social science means looking at the rules that stabilise our expectations about a given action and, at the same time, construct the actor's environment. This reconstruction provides insight into the reasons for action. Society is like a kind of game, the rules of which need to be grasped by the observer.^F For Schmitt the correct description of an action is one that relates it to the agent's reasons (Eyas, 1999, p.15A). This kind of dynamics is absent from natural science. This portrait of an intrinsic difference between the natural and the social realms makes any talk of a 'science of society' in the naturalistic sense extremely problematic. Unlike the mainstream imitations of Friedman's 'positive economics', interpretive social theory may take into account the empathy between those observing society and what is observed. On the one hand, we are able to provide a proper interpretation of the meaning context in which actors are embedded, at least in part, because we share participation in society with them (but see Eyas, 1999, pp.15A-15; Connell, 1999G, p.5F). On the other, although being more aware of this fact than 'value'

^F Schmitt explicitly links his attention to rules in this particular framework to the philosophical influence of Ludwig Wittgenstein, a move that helps us differentiate his view from Weber's.

free' theorists can yield some critical results, we cannot ignore anymore the role played by values in the social action analysed. But can we keep assuming our observation does not interfere with what is observed.

,or even though theory may be serving us, the social scientists, simply as an instrument of explanation, the agents whose behaviour we are trying to explain will be using (the same or another) theory, or proto-theory, to define themselves. So that whether we are trying to validate a theory as self-definition, or establish it as an explanation, we have to be alive to the way that understanding shapes practice, disrupts or facilitates it (Taylor, 1990, p. 100).

In other words, social science is 'reflexive', as it may well be, for instance, Latin to self-fulfilling prophecies, which generate what they foresee (Hollis, 1995, p. 100).

Cespite the fact that Weber and Taylor still take 'explanation' (loosely defined) into account, in IR a sharp divide between explanatory and interpretive theory has often been assumed on both sides of the issue (Barsons, 1997, pp. 100-101). Peeling from the mainstream, Bohane describes the sociological approach in IR and denounces its supporters on the grounds that they need to develop testable theories. For him, a synthesis between this approach and the mainstream will not emerge full-blown, but rather derive from competition and dialogue between these two research programs (1990, p. 100). An important defence of the sociological or interpretive view is provided in the work of Martin Hollis and Steve Smith, where it is juxtaposed to the 'explanatory' mainstream programme. Explanation and understanding, they say, are two valid modes of inquiry which do not go well together (1997, pp. 100-101). In an individual statement at the end of the book, Smith expands on the difference such distinction makes for interpretive IR theory.

I do not accept the idea that we can construct a neutral theory, valid across time and space, that allows us to predict in the same way as occurs in the natural sciences. I do not see that as only a matter of complexity, but as a fundamental feature of the social sciences (Hollis & Smith, 1997, p. 100).

One of the points of contrast here is the role theory should play in social science. Ian Murdoch voices the postpositivist view that in social life data are not fully objectifiable, observers cannot be fully autonomous of the subject under study, and social relationships cannot be separated into discrete 'causes' and 'effects'. By implication, the purpose of theorising is not to identify and test hypotheses about lawful regularities, but to interpret or understand how social meaning and power produce

the apparent stability in the social world (Siddons, p. 10). These differences matter for those defending IR theory as a social *science*, even in the interpretive fashion. Richard Bebbow (1995, pp. 5-6) further expresses the postpositivist IR objection to the narrow understanding of science as a form of inference whose ultimate goal is predictive theories. Social theorising, however, does not necessarily imply any willingness to throw the baby of science away with the bathwater of positivism. That is, 'science' does not equal 'positivism' and does not require strict adherence to the mainstream's view of theory, and in our reflection of mainstream ways of theorising we need not dismiss the notion of a *social science* (E. J. Duvivier, 1995, p. 718). It is with that differentiation in mind that critical theorists and some (but not all) constructivists have produced a new platform for social scientific IR along postpositivist lines (Lotz, 1995). In doing so, they purport to take into account both explanation and interpretation, reflexivity and non-mainstream scientific rigour.

Central to this approach to theorising is the redefinition of the explanatory role of social science. Among theorists in IR adhering to a non-positivist position are some constructivists like Emanuel Adler and Alexander Wendt.

Constructivists have followed third Debate theorists in reflecting the possibility and desirability of formulating law-like generalisations that would justify a meaningfully positivist science of international politics. Yet they have not shied away from offering more contingent generalisations about aspects of world politics. Drawing on their analyses of historical processes, cultural practices, intersubjective meanings and norm formulation they have proffered generalisations about the nature and dynamics of international change, institutional development and moral community (Eric D. Reusmit, 1995, pp. 5-6).

Adler (1995) argues that most constructivists have attempted to re-define the status of social science in IR by occupying a 'middle-ground' between mainstream theory and 'relativism'. He points out the need to unpack the coherent constructivist methodological base that suggests a practical alternative to imitating the physical sciences (1995, p. 9). Commentators and critics, even on the constructivist side, differ on the value of such strategy, but it has certainly called theorists to re-think IR as a social science (Soulé, 1996). Wendt's notion of theory, borrowed from elements of scientific realism, has played a significant role in this process. He notes that one of the essential features of theories seems to be that they 'explain' (1995, p. 10). There are, in his view, two kinds of explanation. Under the influence of logical positivism and falsificationism, mainstream IR has subscribed to the 'covering law' way of theorising. A causal theory,

in this sense, connects several aspects of its subject-matter via a sequence or chain of events to a 'billiard-ball' approach. Wendt's preferred way of theorising, by contrast, attempts to account for the properties of things by reference to the structures in virtue of which they exist. This is what he calls 'constitutive theory', an argument that addresses questions like 'what is *P*' and 'how is *P* possible' (1990, pp. 1-10). While not all constructivists defend a scientific enterprise in IR, the middle-ground presentation of Adler and Wendt has gained acceptance into the core of the discipline (see 2000: 10-11). Their accounts of the 'constitution' of international phenomena in interpretive terms have generated fertile research programmes.

Critical theorists in the general sense are also concerned with the constitution of world politics, but add a clearer transformational and emancipatory aim for IR as a social science (Ashley, 1990, p. 59). Critical theory questions the dominant world order by taking a reflective stance on the framework of this order. By doing so it also questions the origins and legitimacy of political and social institutions and the way they change over time (Griffiths, 1996; Gallagher, 1996; Roach, 1996, p. 1). Here we see a constitutive component. There is also a 'reflective' element whereby critical theory re-opens assumptions that have grounded our political thought (O) by questioning the starting point of thinking politically (Hinsley, 1999, p. 5). Critical theorist Andrew Eynon (1990, pp. 5-9) defends a similar position, but adds a number of important details. Like most critical theorists, he argues that critical theory reflects pre-existing social purposes and interests, collapsing the subject-object distinction and denies the immutability of existing structures.⁶ Eynon (1995) is more explicit on the normative side, in that he believes it is the critical theorist's aim to clearly delineate alternatives in a substantial way. Working in close dialogue with political theory, he stresses the need to think about the emancipatory aim of producing new forms of political community which break with unjustified exclusion (1990, p. 5). In Sened's words, this means freeing people, as individuals and groups, from the social, physical, economic, political and other constraints which stop them carrying out what they would freely choose to do (1990, p. 10). In IR, critical theory's subject-matter is, therefore, two-fold. On the one hand, it purports to look at how the current order has been constituted, with the aim to denounce its patterns of harm and exclusion. On the other, it also investigates the possibility of alternative arrangements

⁶ This unclear connection with the subject-object issue is perhaps an allusion to the fact that we are both students and masters of social reality.

with a view to emancipatory transformation (see Hutchings, 1999, pp. 175–176).

Another group of scholars acknowledge some of the insights of the 'interpretive turn' in social theory but maintain an explanatory and scientific framework. They overlap with some critical theorists and constructivists, but can be identified as a separate group sharing a view of theory influenced by scientific realism (see Haslar, 1999). In other words, the main task of science for them is to explain 'facts' in terms of more fundamental structures (Hutchins, 1999, p. 9). Notice the need to go beyond the surface of phenomena to beyond the empiricist view advocated by the mainstream and uncover 'fundamental structures' in terms of their constitution. This implies, in turn, that a mere 'covering law' and 'billiard ball' account of events cannot fully explain, due to the exclusive focus on observables and assumption of discrete 'causes and effects' (Edström & Linstead, 1997, p. 10). Instead, an alternative notion of explanation is required.

Thus, if asked *why* something occurs, we must show *how* some event or change brings about a new state of affairs, by describing the way in which the structures and mechanisms that are present respond to the initial change. To do this, it is necessary to discover *what* the entities involved are to discover their natures or essences (Cartwright, 1993, p. 10).

(Mechanisms in this sense are frequently occurring and easily recognisable causal patterns that are triggered under generally unknown conditions or with indeterminate consequences (List, 2000, p. 17). The reference to a more nuanced view of explanation as 'causal analysis', referring to structures and mechanisms, has been stressed by scientific realists in IR, like (Ilina & Golin 2007). Scientific explanation is about providing deep understanding of the processes and objects around us, says (Ilina (2000, p. 17). Due to the indeterminate character of causal mechanisms, prediction cannot be the primary aim of science (Ilina, 1999). The theory for 2007 is suggestive of the elements we deem important to the explanation of any given event (Ilina, 2000, p. 17). Such elements are not restricted to material factors. Ideas, meanings, concepts, etc. employed by agents, to the extent that they matter, are part of a causal complex and, hence may be mechanisms (Golin, 2007, p. 17). Ilina concurs, opening up the possibility of integrating interpretive accounts into a scientific realist approach in IR. In order to be able to explain things in this sense, scientific realists broaden the notion of 'causation' and make reference to a variety of things, actions, processes, structures or conditions that we can talk of as being responsible for directing

outcomes, actions, states of affairs, events or changes (M 5//, 5//@, pp..5'.).

IR theory as discourse and practice

A particular reading of Kuhn's 'paradigmatism' has been a key source for the postpositivist view of theory. Reference is made to 'incommensurability' and some of the relativistic implications of Kuhn's account (Mith, 1999, 5//F). For Kuhn, theories are never isolated, but deeply embedded in worldviews or wider intellectual and cultural frameworks. Adhesion to a paradigm is not fully explained in rational terms. By implication, ultimately the dominance of a paradigm is a matter of convention, at least within the scientific community, rather than its intrinsic merits. Rival paradigms cannot be compared in the strictest sense, even if at first sight they may attempt to deal with similar phenomena. Scientific observation, description, explanation and even validation are never neutral, and are deeply bound to paradigmatic understandings. If that is indeed the case, we are actually speaking of paradigms as *constituting* different worlds (see Beufeld, 1999). Steve Mith is a key IR theorist advancing this portrait of theories as a disciplinary narrative (5//, pp...5%)

I think that the theories (O) are like different coloured lenses% if you put one of them in front of your eyes, you will see things differently. Some aspects of the world will look the same in some senses, for example shapes, but many other features, such as light and shade of colour, will look very different, so different in fact that they seem to show alternative worlds (O). The theories we use cannot simply be combined together so as to add up to different views of the same world of international relations; instead, they actually *see* different worlds.

In short, 'paradigm' implies incommensurability (Caddow, 5//9, p.=). Although perhaps not clearly intended, this view of paradigms as constituting different worlds has a radically relativistic potential, which should lead to the 'peaceful coexistence' of a plurality of theories. At least this is how several non-mainstream IR theorists sympathetic to Kuhn's view interpret his claims (see Mith, 1999 for a critique). Of course, this selective reading ignores the other side of Kuhn's paradigmatism: the need to erase this plurality, agree on a single paradigm and 'normalise' IR, turning it into a puzzle-solving science.

An outcome of the 'postpositivist turn' in social science is that it looks very different from mainstream inquiry. The theory now consists, in Clifford Geertz's (1974, pp..95.) words, of blurred genres, whereby historical, philosophical, literary and artistic expressions of interpretation become extremely relevant and 'hypothetical'

deductive, axiomatic and formalised material will be less recurring.

The strict separation of theory and data, the 'brute fact idea'; the effort to create a formal vocabulary of analysis purged of all subjective reference, the 'ideal language' idea; and the claim to moral neutrality and the Olympian view, the 'God's truth' idea – none of these can prosper when explanation comes to be regarded as a matter of connecting action to its sense rather than behavior to its determinants (Zeerdt, 1990, p.10).

Anti-naturalist theorising, therefore, has led the study of society to a revised style of discourse, consisting in the casting of social theory in terms more familiar to gamesters and aestheticians than to plumbers and engineers (Zeerdt, 1990, pp.10-11). Theorising, in the words of a feminist IR scholar, is a more nebulous activity than theory building exercises aiming at constructing law-like statements (Lyvester, 1999, p.10). In fact, as Richard Ashley points out, this blurring of genres to which Zeerdt refers has the effect of turning IR theoretical discourse into practical and prescriptive text whose primary function is not unlike that of speeches, dialogues, treatises, and precepts (1990, p.11). Approaches influenced by the incredulity towards meta-narratives which characterises postmodern orientations (Eyotard, 1999) deny both the desirability and possibility of the 'science' ideal. Paul Heyerabend controversially argued that the only methodological statement that does justice to the history of what we call 'science' is 'anything goes' (1978, p.1). He pronounced his deep concern with the 'authoritarian' implications of a 'paradigm' in 'normal science', a discourse of exclusion dictating the norms of knowledge without necessarily being able to justify itself in absolute terms. In fact, Heyerabend saw little difference between science and pseudo-science, academic and traditional knowledge, research and witchcraft (Berheim, 1978, p.10).^A

Lygmunt Lauman voices some of the postmodernist concerns advancing a similar position. We have a high regard for scientific discourse, says Lauman (1990, p.10), because it allows us to predict and control. When we transpose this view to the study of society, the implications are, to say the least, dehumanising (p.10, original emphasis).

That all such demands amount to is that sociologists should offer advice on how to reduce the freedom of some people so that their choice be confined and their conduct

^A (any commentators unfortunately do not appreciate the irony of Heyerabend's statements to this effect. Upon close examination, it can be seen that 'anything goes' emerges from a demonstration of the absurdity of prescribing *a priori* rules for scientific knowledge, rather than being a prescription itself.

more predictable. A knowledge is wanted of how to transform the people in "question from *subjects* of their own action into *objects* of other people's actions.

Reflexivity means that theory as such, or 'scientific' discourse in general, not only 'explains' or 'understands' the world, but also orders it. Perhaps, even reproducing its own hierarchies and exclusions, reducing it to 'predictability' and 'controllability'. Besides, theory and its perverse effects operate in conjunction as an exclusionary constellation of discourses. Theory establishes 'orthodoxy' and condemns 'heresy'. What is 'heretical', being condemned for rational, 'scientific' reasons, is condemned not only in thought (Lauman, 1997, pp.55-7). Since theory constructs the world via exclusion of alternative ways of 'ordering' the world, to theorise is to affirm one kind of social order against alternative narratives and proposals (cf. Day, 1994). Social theory is a *normative* project, although that agenda has been hidden in various "quests for legitimacy and objectivity" (Grosz, 1996, p. 10, original emphasis). A 'scientific' attitude to the use of theory, in this sense, is not only problematic (for if theory is world-producing, then which world would be theory-testing?), it is also undesirable. It excludes alternative interpretations and crystallises such exclusionary patterns in its construction of order. Theory, in Bina Baym's words, is a form of policing (1996, p. 10). "Upon consideration, Habermas's 'anything goes', taken literally, would appear to be the only viable way." But notice, however, that this sceptical attitude to 'theory' does not necessarily imply a rejection of our 'working definition' of theory at its basic level. In fact, it assumes it, and then condemns it as untenable and undesirable.

In IR, a number of theorists of 'poststructuralist' leanings interact with these ideas. An early textbook in this line makes reference to 'the world-making nature of theory, of theory as everyday political practice' (George, 1996, p. 1). A more recent text further develops the point:

IR theory makes organising generalisations about international politics. IR theory is a collection of stories about the world of international politics. And in telling stories about international politics, IR theory doesn't just present what is going on in the world out there. *IR theory also imposes its own vision of what the world out there looks like* (Geyer, 1999, p. 5, emphasis added).

For the poststructuralist, says David Campbell, there is a reorientation that places theory at the centre of analysis, rather than keeping it as simply a tool for analysis.

. = A close reading of Habermas's works would yield a much less relativistic impression. But, then, postmodernists would probably dispute such exclusionary interpretation.

(5/./, p.55>). Due to the blurring of genres, although there is some sense in which theoretical discourse may be discerned from other narratives, IR theory may also be found in unexpected places like classic IR texts, classrooms, and in more popular sites of culture like film, literature, art, and television (Geber, 5//9, p.55=). (Most often, though, poststructuralists assume some sort of distinction between the 'textual genre' connected to our 'basic definition' of theory, most often found in academic works, and other sorts of 'discourse' which also construct world politics. IR theory, according to Campbell (1990, p.1) for example, is one instance (presumably among many) of cultural practices that serve to discipline ambiguity. It is, therefore, only in a loose sense that 'international theory is everywhere', to parody Martin Wright. In Araya Qalew's words, when we theorise, we are global actors, and those who we normally consider to be global actors are, in turn, also theorists (1991, p.100). What are the implications of this view of theory in research? Campbell's work suggests a possible answer. He defends a logic of interpretation, as opposed to a logic of explanation. His preferred form of interpretivism is one which acknowledges the improbability of cataloguing, calculating, and specifying the 'real causes', and concerns itself instead with considering the manifest political consequences of adopting one mode of representation over another (1990, p.10). The evaluation of theory is also affected, as Jim George (1991, p.10) indicates:

This position, in simple terms, is that the world is always an interpreted 'thing', and it is always interpreted in conditions of disagreement and conflict to one degree or another. Consequently, and for all our attempts to construct scientific means of solving this problem, there can be no common body of observational or tested data that we can turn to for a neutral, objective knowledge of the world.

Rather than 'police thought', then, this alternative view of theory in IR defends 'openness' and would evaluate discourse against discourse itself. Critical criteria can be followed: how exclusionary of alternative 'voices' is the discourse (Ashley, 1991) and here are also literary criteria. If seen as *text*, what about, for example, its 'aesthetic' properties (Ashley, 1991A, pp.15/A) and finally here (but the list goes on), there are political criteria which could be employed (Qalew's, 1991, p.100). That is the discourse's potential for violence: that is *our* role in theorising: constructing the world. Besides the postmodern or poststructuralist critique of theory, there are other reflections of the project of a theoretical social science. A popular one, gaining relevance

> But how can one reflect 'causes' and remain concerned with 'manifest political consequences'?

in IR, is the defence of 'social science that matters', understood as practical, intellectual activity aimed at clarifying the problems, risks, and possibilities we face as humans and societies, and at contributing to social and political praxis (Lyberg, 5//., p.6). Assuming a distinction between the 'natural' and the 'social' worlds, it is argued that the study of social phenomena is not, never has been, and probably never can be, scientific in the conventional meaning of the word 'science' and that it is therefore not meaningful to speak of 'theory' in the study of social phenomena, at least not in the sense that 'theory' is used in natural science (Lyberg, 5//., p.5A). This is due to the fact that, by definition, a scientific theory possesses a general character, aims to be content-independent, predictive, universal and must have rules. But, then, society unfolds in terms which are content-dependent and cannot be reduced to rules, even in an interpretive sense. So (in this view) it is content, and not some sort of meaning-rule, that determines what counts as a social object of analysis or not (Lyberg, 5//., pp.49-55; GA).[@] Therefore, it would appear that social theory in the strict sense of a 'content-independent framework' is not a tenable project (see Bourdieu, 5///). Trying to execute such a project, in fact, has distorted the relevance and role of studying society. Hence the need for a 'social science that matters'.

In this scenario, the purpose of social science is not to develop theory, but to contribute to society's practical rationality in elucidating where we are, where we want to go, and what is desirable according to diverse sets of values and interests. The goal (O) becomes one of contributing to society's capacity for value-rational deliberation and action (Lyberg, 5//., p.=>).

Social science 'that matters', in short, goes beyond both analytical, scientific knowledge and also technical knowledge or 'know-how' and involves judgments and decisions made in the manner of a virtuoso social and political actor (Lyberg, 5//., p.5).

In IR there is an increasing body of reflectivist work that borrows from this literature. It dismisses the scientific and explanatory aspirations of the other approaches while at the same time avoiding a full commitment to 'postmodernism' and keeping a 'constitutive' focus. Brown (5/.5, p.665) points out the rejection of the mainstream view of theory, which in this case is not to be understood in neo-positivist terms as a set of causal laws or 'if-then' propositions inling independent, intervening and

[@] And then we can "question" is *this* a rule? If so, then the argument collapses (*reductio ad absurdum*). If not, then at least in some cases theory and social science, denied in the claim, would be feasible (*non sequitur*).

dependent variables. Camiano de Melice and Francescò Albino highlight the 'reflexive' component, arguing that the more informal kind of theory emerging in this context is a set of practices in itself, which constructs and delimits what is possible and/or impossible in international relations. When theorists speak of world politics as experts, they are part of this broader 'practice'. IR does not simply gaze at 'it' from a safe distance (2005, p. 67). Vincent Couliot and Emanuel Adler (2003, pp. 10; 2004), who have contributed much to the advancement of this way of understanding our discipline, define research along these lines simply as sense making and situatedness allowing for interaction in IR between the several paradigms, thereby distancing themselves from the stricter 'incommensurability thesis'. While the close interaction with the literature of 'social science that matters' is recent, similar themes can be identified in the contributions of anti-naturalist constructivists like Nicholas Ruff and Friedrich Kratochwil, among others (see Jägle, 2003; Qehfuss, 2005, pp. 5A/G). Ruff (2009, pp. 10.G) is very critical of any fixation with theory as a holy grail capable of providing some kind of transcendental knowledge of ensembles of human practices coherent enough to be self-contained. Kratochwil agrees and takes the argument a step further, adding a 'pragmatic' direction. This is evident in his view of theories as guides for understanding social reality and for directing action (2003, p. A9). Theory as such, taken in the strictest sense, is not necessarily the core of scientific research. Perhaps science is best conceived not as a theory-driven enterprise but as a practice among a set of persons who share certain techniques (Kratochwil, 2003b, p. 5). Given this picture of IR research, how should we proceed in evaluating knowledge claims? Kratochwil (2003a, p. F), the presumption of an overarching set of criteria to judge theory from a 'transcendental' perspective is off the table. It is an 'unrealisable plan'. However, Heyerabend's 'anything goes' is also out of question. The 'truth' of theory, in light of this 'pragmatic' approach, has become a procedural notion of rule-following according to community practices (Kratochwil, 2003, p. 5).

Questions on theory and IR

I close this survey of IR views on the definition, elements, uses and validation of theory with some comments. Table ... (below) provides a brief summary of the commonalities and differences across the views surveyed in this chapter. I will expand on the commonalities. After that, I contrast my reading to alternative accounts of theoretical diversity in IR. This is followed by a series of implications and questions leading to the

ne\$ chapter and the rest of the thesis.

Table 4/4 5 6ie/s of theory in IR

Group of theorists	Defining theory	Elements	Uses	Valuation
Classical	Empirical theory in the 'basic definition'. International theory as a different genre, a branch of political theory.	Concepts; propositions; logical connections; normative statements; historical narrative; political commentary.	Provide practical wisdom; clarify moral issues; organise experience; speak 'truth to power'.	Historical accuracy; internal consistency and coherence with wider normative philosophical frameworks; clarification of concrete situations.
Mainstream	Empirical theory in the 'basic definition'. Theory may be part of a research programme or a paradigm.	Concepts; variables; propositions (incl. hypotheses); logical and 'causal' connections; prescriptions.	Explain via causal inference; predict; organise experience; provide policy-relevant prescription.	Consistency; testability; usefulness of assumptions; verification: falsification; predictive power.
Normative	Empirical theory in the 'basic definition'. Normative theory as a different type. Both go often together.	Concepts; propositions; logical connections; normative statements; historical narrative; exposition of tests.	Provide practical wisdom; clarify moral issues; organise experience; articulate "quest for justice; interpret and adapt 'canonical' theories.	Concern for justice; internal consistency and coherence with wider normative philosophical frameworks; clarification of concrete situations.
Critical	Empirical theory in the 'basic definition'. Theory embedded in a social context. Theory also impacts social context.	Concepts; propositions; logical connections; normative statements; historical account; defence of alternative order.	Social critique; theoretical critique; account for status "quo and identify its injustice; analyse its historical constitution; challenge status "quo in search for emancipation.	Historical accuracy; injustice identified; 'Abstracted empiricism' criticised; emancipatory potential; concrete proposals.
Postpositivist science	Empirical theory in the 'basic definition'. Theory embedded in a social context. Theory helps shape social context. (may be part of a paradigm).	Concepts; propositions; logical connections; constitutive account; categories (not 'variables').	Provide an interpretive account of the constitution of order and change in social reality; understand action in a framework of meaning.	Cave to reflexivity, falsification or verification not tenable in the traditional fashion. Valuation depends on the kind of theory (causation: constitution)
Postpositivist non-science	Empirical theory in the 'basic definition'. Theory embedded in a social context.	Concepts; propositions; logical connections; constitutive	Provide practical wisdom; organise experience; provide an interpretive account of the	'Abstracted empiricism' criticised; theory better evaluated, as a practice, by

	&theory helps shape social conte\$ts. (ay be part of a paradigm or worldview.	account; categories; normative statements; reflection on how own theorising affects the world.	constitution of order and change in social reality; understand action in a framework! of meaning.	e\$pers doing scholarship and by its use to the community of political practitioners.
3ostmodern'related	Ac!nowledges empirical theory in the 'basic definition'. -owever, any 'label' or 'concept' is inherently problematic. cholarship should have various shapes and forms.	6oncepts; propositions; logical connections; several types of discourse and critical reflections on narratives of different genres. 6ultural criti"ue.	Against other theories% they 'police' thought and e\$clude discourses. +wn theories% social criti"ue; theoretical criti"ue; provide 'a voice'; deconstruction.	Aesthetics; form criticism and deconstruction; criti"ue of normative and actual implications of theory; potential to 'voice' e\$cluded discourses.

ource% +wn elaboration.

In the beginning of the chapter I set forth a 'basic definition' of theory to which all the positions represented in the table refer in one way or another. &theory is 'systematic discourse about a certain sub9ect'matter'. &he definition, of course, does not mean much unless it is further e\$plained (lunge, .99@a). &his is what each of the views of theory in IR have done throughout the development of the discipline.)et, at the same time, the 'basic definition' tells us a great deal about the nature of these debates. &theory is discourse. It can be e\$pressed in different ways, but loosely spea'ing, it must consist in at least a set of connected propositions (lunge, .99=, pp../@'..G). &theory is systematic. &hese propositions are connected in many forms, and need not be stated e\$clusively with a formal language, as some 'syntacticists' initially believed would be possible. In fact, demanding too much rigour in the presentation of theories would probably 'erase' most IR research from the category of 'theory' (see Rice, 5///, p.5G/ff). &theory is about a sub9ect'matter. 'Aboutness' can be a disputed notion. &here are many ways of theorising the same referent. (oreover, 9ust what 'about' means itself is a matter better left for philosophers in the field of semantics (lunge, .9>5). &he 'sub9ect'matter' is also a matter of deeper philosophical debate. As we have seen, there is much disagreement in philosophy and social science on the status of the things we theorise about. Are they mind'ndependent8 Are they always observable8 Are they useful figments of the imagination8 Bevertheless, all positions agree that theories *do* have a sub9ect'matter. o, then% theory is an argument, consisting at least in interrelated propositions, about something. *p to this point, all views of IR surveyed in this chapter

can agree.

There are also commonalities in terms of the *type* of disagreement across these views. First, the 'structure' of the claims is similar. Most of the positions on theory in IR attempt to flesh out not only this 'basic definition' or 'common denominator', but also the elements, uses and evaluation of theory. These three issues go together in all the positions mentioned here (see Hoover, 1996, pp. 10-11). A certain formulation is articulated about the 'way' in which one should theorise. The elements of a theory will vary according to the way of theorising. This mode also impacts what one assumes to be the use of theory. Much use informs a notion of what makes a proper theory, or how one should evaluate it (Attin, 1995, p. 10). Another commonality here is the 'origin' of the claims. All the positions reviewed in this chapter borrow either directly or by analogy from other disciplines. IR scholars tend to live in an echo chamber (Qaraol, 1995, p. 10). The philosophy of science, certainly due to its focus on theories and ways of inquiry, has been a major source of ideas. A further source is the general literature on social scientific research, and the more critical literature on scholarship about society that reflects stronger claims to science. There is yet another commonality in that all these 'borrowings' involve some recontextualisation and 'blending' in order to accommodate the borrowed material to the new field of IR. As a result, we can see thus the formation of a 'discursive layer' in IR preoccupied with the notion of theory, what to do with it, how to tell its story in terms of the 'growth of knowledge', how to judge it in terms of the context to which it refers, or in which it emerges. This 'discursive layer' is not the mere application of 'eternal' tests to our discipline. It is also a 'bridging' mechanism, hybridising them with our disciplinary views.

Only by investigating how these notions of theory in IR shape themselves in dialogue and tension with their 'origins' can we better understand this 'discursive layer'. This is important for several reasons, not least because it provides a way of criticising our disciplinary 'self-images' about where the divides are and where there is room for cross-theoretical interaction (Eale, 1995). And, for example, the dichotomies we often find in the literature on the types of IR theory% 'normative : positive', 'critical : problem-solving' and 'explanatory : interpretive'. Normative thought and positive theory often coexist and acknowledge each other's use and relevance. Traditionalists like Morgenthau, Hall and Martin Wright certainly recognised alternative ways of theorising IR, defined as 'a system of empirically verifiable, general truths, sought for their own sake' (Morgenthau, 1994, pp. 10-11). Hall (1995, p. 10) provides a similar definition of

IR theory as a body of general propositions that may be put forward about relations among states, or more generally about world politics. Despite Wright's (1991) deliberate focus on common sense, practical and philosophical knowledges while discussing so-called 'traditions' in 'international theory', he has at least once provided his own definition of empirical IR theory% some conceptual system which offers a unified explanation of international phenomena. Speaking on behalf of contemporary normative theorists, Brown (1995, p.5) defines empirical IR theories as sets of interconnected law-like statements which are then embedded into wider intellectual frameworks. A number of mainstream scholars, in turn, known for their narrower view of 'positive' theorising, in fact did not reflect normative and traditionalist thinking either. Waltz (1979, p.1) himself acknowledges a different use for the term 'theory' in the traditional sense of philosophical interpretation, a definition in line with his own earlier work (Waltz, 1979). The distinction between normative and positive theory, therefore, is not as intrinsic to theorising as it would appear to be.

Another questionable dichotomy is that between 'critical' and 'problem-solving' theory. Robert Ross (1991) defends the former and contrasts it to the latter, claiming it tends to reiterate the *status quo*, focusing in incremental 'micro' fixes rather than deep change. The problem with this sharp distinction is that, in and by itself, it does not tell us whether a specific theory is critical or not. (Many critical theorists influenced by Marxist political economy (including Ross himself) are very sceptical about the current economic system. They theorise from a critical perspective in order to change it. However, a free-market libertarian like Hans-Hermann Hoppe (1991) does exactly the same, questioning how we got to the current state of affairs, how prevailing ideology serves the purpose of 'big government' and what we need to do in order to transform the status quo into an anarcho-capitalist system. The main task of contemporary social science, says classical liberal Eudwig von Mises (1999, p.6), is to defy the taboo by which the established doctrines seek to protect their fallacies and errors against criticism. It seems that this way of portraying the distinction between critical and problem-solving theory is confusing since the content of the term critical is dependent on a political context (Wright, 1991, p.5).⁹ It is confusing because two theories can be equally 'critical' in this definition on exactly the same issue and move exactly in the opposite direction. We may still draw a line between more and less

⁹ Critical theory is not the monopoly of Marxist-oriented scholars, and it is fair to say that most normative writers, regardless of buying the Marxist diagnosis or not, may also want to alter world society in one way or another (Williams, 1995, p.5).

conservative theories, but a sharp dichotomy along these lines will not do, if it avoids accounting for substantive claims.

While this 'divide' in the discipline has attracted some attention, the 'explanation vs. understanding' dyad seems to be more widely adopted, especially in some of its alternative manifestations of 'positivist vs. postpositivist' and 'science vs. practice'. For proponents of the sharp divide, this boils down to a difference over what the social world is like. They are of the opinion that explanatory and interpretive theories are not combinable so as to form one overarching theory of the social world (Smith, 199A, p.5). Social scientific attempts to combine both are condemned as unrealistic plans and delusional projects (Kratohvil, 1956, p...). Theory as merely a tool is rejected, and theory as everyday practice is stressed (Qalewski, 199=, pp.FGA=). Of course, the 'unity of science' view inspired by positivism and defended in the mainstream is a way of eliminating the dichotomy at the cost of erasing interpretive approaches altogether from 'social science' (Hernoff, 1957, p.F@). But the sharp dichotomy fails to do justice to a non-positivist alternative, advanced in relevant claims by scientific realists (Atomlin D Wright, 1957, p.5). Kuhl (1958, p.5) points out that many interpretivists reject the ideal of a social science based on causal analysis because they equate 'social science' and 'causal analysis' to 'mainstream social science' and empiricism. Rather than distinguishing between 'causes of' and 'reasons for' action, scientific realists can recognise reasons as causal in the sense that they 'produce' outcomes (Kuhl, 1958, p.5). In a similar way, Golin Wright (195=, p.5) acknowledges that to some extent theoretical discourse may construct social phenomena, but also maintains that a line should be drawn. After all, if the agents engaged in their activities and concepts they have of those activities are exhaustive of the social world, it is difficult to see what role social scientists might play, other than to mystify, through a technical language, that which social agents already know. In fact, Stefano Zucchini (1970) has shown that some of the anti-naturalist interpretive work can be re-theorised in terms of 'causal mechanisms', and suggested, with Wright (Zucchini, 1970; Wright, 1957), that interpretivists also tacitly assume scientific realist principles in their practice of scholarship. Since they value taking practice seriously in its own context, then the recurrence of this tacit assumption suggests they should take scientific realism more seriously.

After the survey in this chapter, a more nuanced view of the theoretical divide in IR emerges. There are more points of contact than the dichotomous 'self-images' in the

discipline would allow. Because a number of the alleged great debates in our field has been portrayed in function of these binaries, we can safely add this contribution to a more accurate account of the disciplinary divide in IR. These are by no means the only 'organising' schemes suggested by theorists to account for theoretical diversity. Some prefer to distinguish theoretical discourse according to substantive claims about the international system (AttinT, 5/..; Zabriel, .99G; Woods, .99=). Others focus on the outreach of the theoretical scheme (Jinger, .9=.; Waltz, .9>9, pp.=/'>@). But the classification such as the one pursued here, in terms of 'ways of theorising' IR, has been increasingly favoured (Urgensen, 5/./, pp.5/'). The findings in this chapter strengthen the categorisation proposed by Kuhl and Wright (5//>), but suggest we should perhaps go beyond that. The main types of IR theory, according to them, are 'explanatory', 'critical', 'normative', 'constitutive', and 'theory as a lens'. I would add that we need to bear in mind that these are 'types', i.e., not exclusive, but capture the most salient claims in each sort of theoretical discourse. I would also add that theory is often portrayed as embedded within a wider framework, be it a paradigm or research programme. With these authors, I would argue that "questions of theory appraisal, objectivity and impact of theory on practice (or vice versa) depend on the philosophical background that shapes each of these ways of theorising.

Another level of discourse

At this stage we may draw nearer to the question posed by this thesis. A number of themes to which this chapter alludes raise issues about another level of discourse in IR: discourse about theory and the discipline themselves, and not about world politics *per se*. First, we have seen that in IR several views of theory, its main elements, uses and evaluation are in part constituted by pre-existing views in philosophy and general social science. Secondly, it has been pointed out that, on several accounts, the material borrowed from these other fields is often adjusted to suit IR scholars' concerns. Thirdly, some of the issues addressed at this level of discourse deal with how to judge theoretical work and how to theorise. The scientific status of IR, the political and social impact of theory and the relation between theory and its subject-matter are also discussed at this level. Fourthly, one could add critical concerns to the list as well. For example, narratives about theory tend to generate accounts of the historical formation of IR as a discipline, such as the great debates story with the dichotomies upon which it relies. This other level of discourse would provide space to question the 'self-images' of the

discipline. It would also provide critical space to go beyond its role as a 'bridge' between philosophy, other disciplines and IR and also criticise such attempts to 'bridge' and 'hybridise' material across fields. &hese issues pertain not to the study of world politics, but rather to the study of *scholarship* about world politics. &hey are located at the 'bridge' level, the discursive layer placed between one theory and another. &hey are, in this sense, "uestions of a *metatheoretical* nature.

Chapter &

'hat is Metatheory(

The "uestions raised in the previous chapter about the nature of the discussion on theory in IR led us to the assertion that they belong to a specific 'discursive layer' in the discipline. While many of these debates are philosophical in nature, they are also shaped by a specific disciplinary context. For this reason perhaps we should consider them in their own terms, instead of hastily reducing them to philosophy. Despite being raised and debated in IR, they are not so much "uestions about world politics as they are about the issue of theory itself 0 in this case, the theory of international relations. In other words, they are problems of *metatheory*. Let a theory "uite simply means theoretical reflections *on* theory M and L promises to enhance our understanding of the nature of IR theory M (4Urgensen, 5/./, p..A).

In this chapter, I further e\$pan the discussion by loo!ing at the definition and types of metatheory. I begin by reminding the reader of our preliminary definition of 'metatheory' as 'theory of theory' or 'systematic discourse about theory'. After highlighting some of the implications of this definition, I note the use of parallel understandings in general social science and two selected cognate fields to IR 0 economics and political science. Following that, I unpack the preliminary definition in a typology of metatheoretical research that accounts not only for the e\$amples provided to illustrate each type, but also for similar attempts to classify metatheorising in social science. In this typology, the general roles of metatheoretical research are described according to their most salient features following two criteria. On the one hand, there is the sub9ect'matter of metatheoretical research. In connection, of course, to the centrality of 'theory', other issues are addressed by metatheory. On the other hand, these issues are addressed in different ways 0 metatheoretical discourse will have distinct functions depending on the type of argument. Metatheorising in IR reflects, therefore, a rich combination of elements on both sides of the typology. Having established a definition and typology for metatheory, I proceed to a discussion of metatheoretical research according to the !ey combinations of sub9ect'matter and predominant function using IR

examples. This I do in order to open up a more nuanced understanding of the role of metatheory in the discipline in general terms. With that, part one of the thesis is closed. An exposition and evaluation of what IR scholars themselves claim to be the roles played by metatheoretical research in the field is left to the forthcoming chapters, in the second part of my study.

Theory of theory

The concept of metatheory adopted up until this point defines it as 'theory of theory'. But what are the basic forms in which we can systematically study IR theory? What exactly are the implications of this way of thinking? Let me begin to further explore the issue by showing in this section that the preliminary definition of 'metatheory' adopted so far is consistent with IR usage. In the next section, I turn to parallel views in general social science, economics and political science.

The notion of metatheory as systematic discourse on theory in broad terms is no novelty in the academic discipline of IR (Griffiths, 1996; Gallagher, 2000; Roach, 2001, p. 10). (Griffiths, 1996, pp. 1-2) highlights reflexivity as the core of metatheoretical research. He contrasts IR theory, which treats empirical evidence, with a certain type of theory that addresses the issue of theory itself. And in this sense, metatheorising consists in reflection on the process of theorising. Alexander Wendt (1992, p. 10) poses a distinction between first order theorising in the form of substantive theories and second order or meta-theorising. Roger Pegelow (1993, p. 111) clarifies the meaning of second order inquiry, or research whereby one is examining the examiners of the world rather than the world itself. Perhaps in a more judgmental way, Hedderley (1996, p. 5) contrasts meta-theory, focused on theory, with substantive analysis. Colin Wright (2005, p. 11) draws the line between metatheoretical research and IR theory in terms of a distinction in subject-matter between the nature of inquiry itself, as opposed to the nature of the international system. With (Hedderley, 1996), he further expands on the difference between first and second order inquiry. Meta-theory does not take a specific event, phenomenon, or series of empirical real world practices as its object of analysis, but explores the underlying assumptions on the act of theorising and the practice of empirical research (Griffiths, 2005, p. 1). In short, they say, metatheories are theories about theories.

Another way of thinking of metatheory in IR is to define it as a 'bridging'

discursive layer between the generality of philosophy and the specificity of theories of world politics. This adds an extra dimension to our basic definition, but does not contradict its core. Jørgensen (2007, pp. 17–18), for example, makes reference to levels of reasoning and arranges them from wide to narrow in the following order: philosophy, metatheory, theory and empirical analysis. Metatheory is the level concerned with the study of theory. Interestingly, in IR it does not need to focus solely on international theory, but may also make reference to social theory in general, operating as a kind of 'bridge' or 'link'. And in this sense of a 'bridging mechanism', metatheory adopts a more dynamic and synthesising role, by applying philosophical notions to a specific discipline or simply by joining different theories into one coherent approach. Debates on philosophy of science (Kernoff, 2006a), epistemology (Epstein, 1999), ontology (Wendt, 1999), mode of inquiry (3. & 4. Jackson, 2001b), role of theory as social critique (Beufeld, 1999A) and so on are examples of this 'bridging' element of metatheoretical research in IR. While we may notice an additional dimension here, we need not interpret this feature as detached from our basic definition of metatheory. Jørgensen (2007, p. 59) himself expands on the 'bridging' role played by metatheoretical research with reference to each of these topics and still subordinates it to the notion of 'systematic discourse about theory'. In this case, operating in a dynamic mode as the framework of analytical commitments by means of which it is possible to reflect on existing theories or create new theories. These positions of metatheory, mediating philosophy and theory, or as a 'bridge' between one theory and another are, thus, some of the key features IR scholars ascribe to metatheoretical research.

A final way in which this discursive layer is identified in IR has to do with applications of 'meta-science' or 'science studies' to the discipline. This eclectic field traces the history of disciplines, the dynamics of science as a social institution, and the philosophical basis for scientific knowledge (Kernoff, 1999, p. 1). And here is, for example, an increasing number of works on the sociology of knowledge in connection to the production of IR theory (Acharya, 2001; A. 1. & 2. D. 7aeber, 2009; &urton, 2001; &urton D. 7aire, 2009). And here is also an increasing number of in-depth studies on the immediate settings that shaped certain episodes of our disciplinary history (Zuilhot, 2001c; Robson, 2005; Einlater D. uganami, 2001=). This type of research in IR is less frequently self-identified as metatheoretical, as it tends to focus on additional aspects of theorising, such as social context and disciplinary history. However, one cannot help but notice that such works on the 'contextual' elements of IR theory share the feature of

being 'systematic discourses about IR theory'. It is here, however, that I part ways with those who defend a clear-cut dichotomy between 'first and second-order theory'. (Metatheoretical studies will often either directly or indirectly uncover mechanisms in the 'real world'. Theories are part of the social world, and it should not surprise us that a study of the social world may include analysis of how it shapes theory and how theory shapes it to some extent (Wallerstein, 1986, p. 5). Provided we keep in mind that metatheory is only a specific subset of this kind of research (namely, study that has theory as its key subject-matter) we can thus infer some limited overlap between first and second-order. It is clear that metatheory is generally understood in IR as 'theory of theory'. Before we look at further implications, let me point out that cognate disciplines in social science share with IR a similar view of metatheoretical research.

Metatheory in cognate disciplines

(Arieh Elster, a philosopher of science whose work delineates philosophical analysis as an exercise that cuts across all theoretical disciplines, speaks extensively of metatheory in his contributions to social science. According to him, there are at least two points of entry for metatheoretical research in this context (1990, pp. 1-2). On the one hand, social science includes philosophical material to a great extent. On the other, specific controversies which have emerged in key social scientific fields are inherently philosophical, in that they permeate all theory-based disciplines and point out to issues of a broader nature (e.g., the nature of reality, the validation of knowledge, the role of scholarship). 'Metatheory', in a sense, is part of the philosopher's duty towards this set of specialised fields of knowledge (1990, p. 11). This definition of the term echoes what we have seen so far in the IR literature. A metatheory is a theory about theories. Roland Diekse, another philosopher of science, stipulates that theories of science (including theories about theories) are analogous to scientific or empirical theories. That is, in principle they should have similar structure, consistency and formality (1999, p. 1). Elster, in turn, warns against being too strict in defining what a 'theory about theory' is, considering that what we frequently call 'metatheory' in social science tends to be sets of loosely knit metatheoretical propositions (1990, pp. 15-16). Similar informal views of metatheory as 'theory of theory' abound in the social sciences in general and, more specifically, in fields which display many affinities with IR, such as sociology, economics and political science.

Steven Wallis (1997) has compiled and analysed a number of definitions of

metatheoretical research in social science. One of the most relevant conceptualisations listed in his study corroborate in other disciplines what we have seen for the case of IR. Lakatos (1976), defines communication theorist Robert Braig, his theory about theory. The purpose of metatheory in general has to do with the uncovering and critiquing assumptions of theory (1979, p. 14). Deborah Kingfeld, a health research specialist, fleshes out some of its functions. Analysis and interpretation of theoretical, philosophical, and cognitive perspectives; sources and assumptions; and contexts of scholarship are central among the roles of metatheorising (1978, p. 19). Even in health research, the authors of an authoritative handbook define metatheory as 'a critical exploration of the theoretical frameworks or lenses that have provided direction to research and to researchers'. In this sense, it relates to the analysis of primary studies for the implications of their theoretical orientations (Waterson, & Horne, 1980, Dillings, 1977, p. 95). Social theorist Yang Qiao asserts in an encyclopedia entry that metatheorising 'focuses on the examination of theory and theorising'. He adds that metatheory takes place in virtually all fields of social science (1978, p. 14). In his study of these definitions and parallel statements of social scientists, Wallis (1977, p. 14) concludes that most reflect the 'core' of metatheory and a number of them address only some of its aspects. Lakatos (1976), in his synthesis, is primarily the study of theory. This includes both the development of overarching combinations of theory and theorems for analysis that reveal underlying assumptions about theory and theorising. As seen in the case of IR, while there are other peripheral aspects to it, the key to metatheory is that it revolves around 'theory'.

(Metatheoretical issues are often discussed under the label of 'methodology', although this is not necessarily a universal trait of all disciplines. In economics, for example, the term 'economic methodology' is frequently used interchangeably with 'philosophy of economics'. As Daniel Hausman, a key specialist in this subfield, puts it, the methodology of economics is 'a philosophical questions'. They include issues such as the goal of theorising, the definition of a theory and a model, how they relate to reality, and whether there is an intrinsic distinction between natural and social science (1978, pp. 14). But, then, these are metatheoretical questions. Drawing on a similar set of questions, Glenn Ross, another economic methodologist, comes to the conclusion that methodology, in this particular sense, is simply the theory of theories (1990, pp. 14; 15). Unlike general philosophers of science, who apply philosophical reasoning to specific fields of science from outside, economic methodologists (or metatheorists in

economics) philosophise within the discipline of economics (Ausman, 1990, p.10). This is arguably what accounts for the philosophical eclecticism and hybridisation which one finds in economic methodology (O'Shea, 1997, p.5).^{5/} Rather than being a mere application of philosophy to economics, this sub-field plays the role of a 'bridge' between general philosophical questions and 'mid-range' points concerning that specific discipline.

When the term 'metatheory' is employed with a clearer meaning, it often refers to the application of the philosophy of science or philosophy of social science to a specific field. Political scientists, who use the term more often than economists, are used to thinking in terms of a 'bridging' role for metatheory. David Marsh (1997, p.5.5), for one, connects (but does not limit) the term 'meta-theoretical issues' to the following list of problems: 'the relationships between structure and agency, the material and the ideational and stability and change'. Speaking from a different perspective, Charles Glynn (1990, p.10), another influential scholar, lists possible 'bridges' between 'epistemology, ontology, and logics of explanation' on the one hand and social theory on the other as 'key aspects of metatheory' (see also Glynn, 1990, p.5). Joseph A. DiMaggio (1990, pp.5-15), another political scientist, provides a similar list, adding 'disciplinary' aspects like the assessment of 'different architectures of inquiry' and 'scholarly style'. Sociology, however, is perhaps the one discipline in which these issues have been more thoroughly studied with a clear view of their 'bridging' role between philosophy and social science (see Ziddens, 1999). While some have adopted a more restrictive view of metatheory, applying it merely to 'theory construction' and 'evaluation' (Ager, 1995, pp.1-10), others have conflated the term with what Wright Mills (1949) calls 'grand theory' (e.g. Bowers, 1997, p.9). However, most would still agree with the general definition developed here of metatheory as the 'study of theories' (Wallace, 1995, p.10).

So a great extent, it was George Ritzer's (1990, 1991, 1995) careful work fleshing out the concept and roles of social metatheorising that set the tone of this sub-field in sociology. Choosing Lunge's recommendation of a more flexible understanding of 'metatheory' to accommodate to current usage in social science, Ritzer (1997, p.10) defines metatheorising as 'systematic study' of theoretical material. In sociology, for example, 'a metatheorist is one who studies sociological theories of the social world, while a theorist is one who studies the social world more directly in order to create (or apply) sociological theory'. This avoids the restrictive view of metatheorising

as a heavily formalised endeavour. (Moreover, it also clarifies how metatheory differs from grand theory. While one of the key roles of metatheorising may be that of cross-disciplinary and cross-theoretical synthesis, the 'bridging' role that we should not lose sight of the 'theory-centred' character of metatheoretical research (Ritner, 1996b, p.6). So be sure, grand theories like Marxism may operate as cross-disciplinary metatheorising in a synthetic move (say, by applying an account of social structure to IR theory) or perhaps in the less common use of Marxism as a 'sociology of theoretical knowledge' (ERwy, 1990A).⁵ However, unlike metatheory, Marxism is *primarily* interested in providing an account of the historical formation of society in general. And in this sense, the looser use of 'metatheory' in some circles would blur the distinction. Simple as it may be, Ritner's definition handles this problem. Another issue sorted by his extensive effort to clarify what it means to do social metatheorising refers to the 'direction' of metatheory. We often read about a certain 'normative' drive of metatheoretical research (e.g. Ross, 1997, p.15; Wallace, 1995). Systematic discourse about theory tends to aim for the improvement of theoretical material in one way or another. This has led some IR scholars, for instance, to affirm that metatheory must always provide clear directives on how to theorise or evaluate theory (Chernoff, 1990a; 3. & Jackson, 1990a). But this is not the only possible function of metatheory, or even a necessary one. While Ritner (1990, pp.6A) acknowledges these roles of metatheorising as paramount, his broader definition does justice to both current usage of the term in social science and the broader notion of metatheory as 'theory about theory'. In his own words, the kernel of this kind of research meets the need to better understand social theory (Ritner, 1990, p.). As we can now tell, this view recurs not only in IR but also in parallel disciplines.

Typology of metatheorising

The discussion above clarifies a number of points. First, metatheory has theory as its key subject-matter. Secondly, by implication, metatheory operates as a 'bridge', between philosophy and specialised theory or between two specialised theories. Thirdly, there are many possible ways to actualise this bridging role. Fourthly, some of these ways may go beyond the academic discipline in question. Finally, some of these ways may involve the relation between theory and its subject-matter in that discipline. These five points invite a typology of metatheoretical research highlighting salient features of

5. I owe this point to BoJ Gornago, who mentioned ERwy's contribution to the sociology of knowledge while commenting my earlier work on the use of Imre Lakatos' philosophy of science in IR.

specific claims. In this typology I follow, in part, Ritzer's account of metatheorising.

Ritzer divides metatheorising in three major types, according to their respective end products (1996b, pp. 1-2). The first type results in new social theory. A second type synthesises theories and produces an overarching perspective. A third type, more frequently encountered in the literature, leads to a deeper understanding of theory and may be subdivided into a focus which is mainly 'Internal' or 'External' to the discipline in question; as well as 'Intellectual' or 'Social', depending on whether we look at theory by itself or theory in context. Ritzer's formulation has been employed time and again by other social scientists and has helped shed light on crucial issues in sociology and correlate disciplines (see Qahar, 1999).⁵⁵ For all its merit, this typology still requires some adjustment. One justification is that a key motivation to systematically study theories is to better understand them. In principle, all types of metatheoretical research pursue that goal, broadly conceived. For this reason, we should not restrict it to a mere subtype. Another justification is that the two axes (Internal/External and Intellectual/Social) also apply to the other types of metatheoretical discourse. Nevertheless, strong reasons abound to retain much of Ritzer's effort, adapting rather than rejecting it altogether. Unlike Ritzer's differentiation of 'end products', I make a distinction between a typology of the *subject-matter* of metatheorising (besides theory itself, of course) and *ways* of metatheorising. This allows for a more nuanced understanding of the functions and roles of metatheoretical research in social science whilst at the same time keeping all of Ritzer's core types of metatheory. My departure from Ritzer is clearer in the next section, where I deal with ways of metatheorising in connection to possible subject-matters. In this section, the main point is to arrange metatheoretical research according to subject-matter. Unlike any other typology, the central concern here is to organise the material highlighting the peculiarity of each type. The aim is to increase our understanding of how metatheoretical discourse operates in a discipline.

In this typology, I draw two distinctions related to the main focus of specific instances of metatheoretical research. On the one hand, theories are often derived from other disciplines and fields, or at least make reference to material outside the discipline in which they are primarily formulated. In order to gain insight into these moves, metatheoretical research may actualise its 'bridging' character by looking beyond its

⁵⁵ In this thesis I shall repeatedly apply the capitalised terms (Internal, Intellectual, Contextual and External) to refer to elements in the typology.

main discipline of focus (Alliday, 1996, p.9). Thus we may speak of study of theory with core references to aspects which are 'Internal' or 'External' to the discipline in question (Ritner, 1996, p.5A). On the other hand, however, we must remember that a discipline cannot be fully reduced to its academic output (Lunge, 1996, pp.1-2) and that metatheoretical research might reach beyond a discipline's content (e.g. Solanyi, 1996). Because theorising also operates in a certain context, and studies aspects of the 'real world', metatheoretical discourse may focus not only on the 'Intellectual' side, but also on the 'Contextual' side of a discipline. The former deals with the cognitive structure of the field, whereas the latter emphasises the way research functions in a community in practice (Ritner, 1996, p.5). In a preliminary fashion, then, we may organise metatheoretical research according to its central focus on features which are *Internal* and *External* to the field, as well as *Intellectual* and *Contextual*. In any case, theory is the 'central subject-matter' around which these additional issues revolve. Theory is like the sun in our solar system: metatheoretical research will 'reflect' its 'light', but in a diversity of approaches in any given discipline.

Table 14.5 Metatheorising according to focus' with examples

	Internal	External
Intellectual	<p><i>Box 1</i></p> <p>Schools of thought IR theories Cross-theoretical debates Assumptions and structure of theory</p>	<p><i>Box 2</i></p> <p>Theorist's immediate circumstances Research funding for specific projects Networks, academic politics Academic prestige in IR</p>
Contextual	<p><i>Box 3</i></p> <p>Borrowings from other disciplines Bridge from philosophy to IR IR and natural/social science Methodological assumptions of theory</p>	<p><i>Box 4</i></p> <p>Impact of general historical context Impact of culture (e.g. British Western) Impact of theory in society IR theory and policy-making</p>

Source: Based on Ritner's typology.

Table 5.. (above) illustrates the possible combinations with generic examples. Each of these types of metatheorising contributes in a particular way to a better understanding of theory in a given academic discipline. Although the general logic applies (either actually or potentially) to any academic discipline, for the purposes of this thesis, I shall refer to examples in IR. Internal: Intellectual metatheorising (*Box 1*)

addresses IR theory primarily in terms of the way it operates in the discipline of IR (including its sub'fields) and places heavy emphasis on logical and cognitive aspects of scholarship. Discussions on schools of thought, refinement of the structure of an IR theory, accounts of cross'theoretical debates and e\$aminations of the underlying structure of a theory are counted among e\$amples of Internal:Intellectual metatheorising. Botice that these instances are all elaborated from the particular perspective of the discipline. In our illustrations, schools of thought would be referred to in the common IR parlance of 'isms' ('6lassical Eiberalism' in IR is very different from '6lassical Eiberalism' in political theory and economics). An IR theory would be refined with the use of critical thin!ing and reference to further IR literature. 6ross'theoretical debates would be described in the discipline's own peculiar narratives (e.g. the 2reat Cebates). &thus, in Internal:Intellectual metatheorising, little reference is made to elements outside the scope of the cognitive aspects of IR.

&he second basic type of metatheoretical research according to sub9ect'matter is #Sternal:Intellectual (*Box II*). It is still primarily interested in the intellectual side of academic production, but loo!s at IR theory with great concern with scholarship underta!en outside the field. It involves Lturning to other academic disciplines for ideas, tools, concepts, theories, and the li!eM and adapting them to IR, or perhaps observing how they have been adapted to IR (Rit?er, 5//., p.59). It is in such 'borrowings' that the 'bridging' role of metatheory becomes clearer, as a bridge from philosophy to IR or from other disciplines to IR. Botice that this does not mean that metatheory only functions as a bridge when it is '#Sternal' in this sense. ,or e\$ample, an 'Internal' bridge would be the 'Beo'Beo' synthesis attempt to combine neorealist and neoliberal institutionalist theories (7aever, .99=, pp.=F'G). An instance of '#Sternal' bridge would be <enneth 7alt?'s (.9>9) adaptation of structuralism to IR. #valuating the coherence of 7alt?'s move in light of sociological structuralism and whether IR gains from it or not is also #Sternal:Intellectual metatheorising (e.g. Ashley, .9@G). *Box I* and *Box II* cover the main types of metatheoretical research focused on !ey intellectual aspects of IR theory. &his !ind of argument provides an in'depth loo! at theory as such, or perhaps, theory in con9unction with other theories and remaining discursive elements of scholarship.

(etatheoretical research that fits into *Box III* and *Box I6* bring the conte\$tual side of theoretical research in IR to the centre. Internal:6onte\$tual metatheorising (*Box III*) highlights additional aspects shaping the development and modification of theoretical discourse in the discipline. &his type of research is predominantly focused on

the specific context of IR. A study of the British Committee on the Theory of International Relations, for example, containing relevant information on the network itself, each member's contributions and how these features shaped English school IR is Internal:Contextual metatheorising (-all, 5//=; Hige??i, 5//A). It tries to connect these communal elements to IR theory in light of the immediate disciplinary setting. #Sternal:Contextual metatheorising (*Box 16*) deals with similar non'intellectual aspects shaping IR theory, but stresses extra-disciplinary factors, assuming they are also relevant to a more complete understanding of the process of formulation, testing and application of IR theory. An illustration is the popular narrative that changes in the structure of the international system (the interaction between Great Powers, general wars, relevant peace settlements, etc.) have somehow shaped the development of IR theory (Gunn, 1999). It is important to remember that this relation between context and theory is not unidirectional. While it is easier to realise that both the discipline of IR and the practice of world politics affect theory, theory in turn can also have an impact in the discipline and, perhaps to a lesser extent, the 'world' of world politics.

Table 5.5 (below) expresses my preferred formalisation of the distinction between these types of metatheorising. Sharp dichotomies between Internal and #Sternal, and Intellectual and Contextual are erased. Instead, both axes are represented in function of these extreme 'ideal types', and specific metatheoretical arguments may be classified in a continuum.

Table 5.5 Metatheorising continuum' according to focus

Internal:Contextual		#Sternal:Contextual	
Internal	Contextual	#Sternal	Contextual
Th _{IR}	Th _{IR}	Th _{IR}	Th _{IR}
Ph	Th _{<}	Th _{IR}	Th _{IR}
		Th _{IR}	Th _{IR}

Source: own elaboration.

In the Internal:#Sternal axis, internal metatheorising may yet allude to #Sternal features, but still retain its predominant focus on internal elements. For instance, even if we are studying the inner logic of Wendt's views on agency and structure and how it is shaped by specific IR debates, we may study Anthony Giddens' original sociological

formulation of the same problem in order to better interpret Wendt's test ($Th_{IR} = Th_{<}$, where Th_{IR} means 'theory in IR' and $Th_{<}$ means 'theory in another discipline'). The same applies to the Intellectual:Contextual axis. Intellectual claims may also resort to contextual elements while remaining in line with the chosen point of entry into metatheoretical research. For instance, awareness of the recently disclosed fact that Adam Watson was responsible as a diplomat for anti-Soviet propaganda may have an impact on how we read his theoretical work on the 'historical tendency' toward hegemonic 'spheres of influence' displayed in several states' systems ($Th \ W \ I$, where Th means 'theory in IR' and I means 'general context').^{5F} The point here is simply that types are not 'pigeonholes' in the stricter sense. The most salient features are highlighted by separating one type for another, but they are not meant to reflect exclusionary binaries.

Metatheoretical research

Internal:Normative and Intellectual:Contextual are combinations of the most salient features of metatheoretical research according to its main subject-matter or *focal point*. In addition to this, we can also analyse ways of metatheorising. There are five main ways of metatheorising: hermeneutical, corrective, evaluative, critical and historical. While no metatheoretical research follows only one of these logics in a pure way, we can often identify which of these 'ways of metatheorising' qualifies a given argument, or set of 'meta-statements' (Lunge, 1999, p. 5A). Table 5.5 (above) contains symbols representing key relations in metatheoretical arguments, which we explore below according to their occurrence in each type and way of metatheorising. Th is IR theory studied by itself, and $Th_{IR} \ W \ Th_{IR}$ relates two or more IR theories. $Th_{IR} = Ph$ relates IR theory and philosophy. $Th_{IR} = Th_{<}$ operates in a similar way, but with reference to theory in a discipline other than IR. $Th = I_{IR}$ connects IR theory to the disciplinary context, and $Th = I$ links IR theory to the broader context. Each relation plays distinctive roles depending on the combination of focal point and way of metatheorising, helping us better understand IR theory.^{5G}

>hermeneutical metatheorising plays the role of interpreting theoretical material.

Internal: Intellectual metatheory of this sort occurs primarily in two forms in IR. It can

^{5F} Lunge and Eittle (5/9, pp. 51-5viii) mention this biographical detail but avoid inferring a strong connection to Watson's arguments and work.

^{5G} These symbols help us locate metatheoretical arguments in the diagram and will be employed subsequently in this thesis. Here, the form 'A W I' does not indicate the logical relation 'iff' (if and only if) between A and I. It merely indicates that a metatheoretical argument is relating A and I, where at least one of these terms is a theory.

look at a certain theory by itself, like textbooks sometimes do. Internal: Intellectual hermeneutical metatheorising also studies the relation between two IR theories to shed light on the meaning of both. For example, students are often required to compare and contrast Walt's 'defensive (neo)realism' to John Mearsheimer's (1993) 'offensive (neo)realism' so as to better understand both approaches. External: Intellectual cases are more frequent. They can relate IR theory to philosophy in an effort to understand the latter's impact on the former. Or they may look at the roots of a certain IR theory in theoretical material found outside the discipline. A philosophically focused work is Kenneth Waltz's (1993) investigation of the origins of IR thinking on sovereignty in political philosophy. An hermeneutical effort connecting IR theory to theories elsewhere, in turn, is Aaron Beck's (1996) study of the relation between Walt's and Wendt's views of social structure and classical anthropological theory. Contextual metatheorising in this interpretive sense, however, is much more frequent than the Intellectual variety for the reason that reading a text with reference to context is a widely practised exegetical norm (see Schleiermacher, 1996). Internal: Contextual research here moves from context to theory. Jim Cumey's (1996) interpretation of the English school in light of archival sources on the meetings of the British Committee is a good illustration% this is Contextual metatheorising, but Internal to the discipline of IR. While the theoretical text still regulates the reading, awareness of an author's situation of writing (what exegetes call *Sit? im @e en*) can often disclose new ways of looking at that text. Even the interpretation of material often taken for granted, such as Martin Wright's (1996) *Power Politics*, can change in light of relevant biographical or contextual information. &his External: Contextual avenue of research I have pursued, discovering connections between hard passages in the text and Wright's Christian pacifist activism, challenging the more conventional reading of that work as a classical realist manifesto (E. 2. Jure, 1996). That all these cases share is their key interpretive function, which qualifies the hermeneutical way.

The second way of metatheorising is *corrective*, in that it helps us adjust and modify theoretical material in order to refine it (see Hendry Dasillos, 1997, p. 10 for a philosophical exploration of this theme). From an Internal: Intellectual perspective, we can focus either on an IR theory in isolation or relate different IR theories. Whether in isolation or in interaction, formalisation of theoretical material basically involves restating the original approach(es) in new modes, and is often pursued as a way of refining it, not necessarily with much success (Lunge, 1999, pp. 91-2). On the

Internal: Intellectual front, refinement is sought with reference to theoretical material from outside. Here we see once more the bridging role of metatheory from philosophy to IR or from theory in another discipline to IR theory, such as Robert Jowett's (1995) attempt to present Waltz's neorealism in game-theoretical terms. External occurrences of corrective metatheorising involve different processes. While most work on the Intellectual side deals with bibliographic analysis, here we face a whole range of techniques employed to refine theory in light of its relation to the 'world' it attempts to explain. The most frequent form in which this occurs is External: Internal, where theory is juxtaposed to events in world politics, and refined in order to better account for them. (any debates on IR theory after the end of the Cold War were intended to improve our models of unipolarity or hegemony and did not necessarily emerge as direct results of theory testing or 'evaluation', but rather as indirect response to this general context. This will suffice to explain and illustrate correction and refinement.

Evaluation, in fact, is our third way of metatheorising. It is often pursued as a form of adjudication of claims to knowledge, but can also go after alternative goals (Tafel, 1995, pp. 5A-B). The need to refer to empirical aspects of scholarship (regardless of one's views on empiricism) means *evaluative* work will be more concentrated on the External side. Still, one can evaluate IR theory in an Intellectual way in terms of coherence, logical adequacy, and so on. Internal: Intellectual evaluation focuses on standards of the IR discipline itself. For example, Starr's smear campaign against 'idealism' or 'utopianism' is an example of internal standard that prevailed for a long time in the field (Kearse, 1995; 3. Wilson, 1995). With reference to material outside the discipline, External: Internal metatheorising either tests IR theory's synthetic efforts to 'bridge' between approaches from philosophy or other disciplines to IR (Coty, 1999) or, alternatively, evaluates IR theory in light of philosophy and these other disciplines. The heavy emphasis on the philosophy of science is a key feature of this way of metatheorising (3. & Jackson, 1995b). However, most instances in this case also point to the External side, with a view to empirically testing the claims of IR theory and/or generally establishing scientific credentials for the discipline, or an approach within it. Scholars are usually inclined to label this kind of exercise 'methodology', but when the philosophy of science is explicitly mentioned as the 'umbrella' that provides a verdict, the metatheoretical character is highlighted (Eale, 1995). Conscientious investigators can't get away with a theory of the phenomenon under study and a simple hope that their methods test the theory. They need two

theories% one of the phenomenon, and another of the process producing their evidenceM (&illy, 5//@, p..9).

#valuative metatheorising is in many respects similar to a fourth type, the *critical* approach. Indeed, they both share the attitude of adjudication, but suggest different and not necessarily exclusive ways of judging theoretical material. Criticalism may be Intellectual:Internal with reference to ideology, normative assumptions and worldviews influencing IR theory (Zabriel, .99G; Griffiths, 5/., pp.5'>). Critical metatheorising is often contextual, in that it looks at the effects of social context on theory and vice-versa. Contextual:Internal criticism analysing the impact of IR theory on the discipline can be illustrated by studies on the politics of the discipline. Cynthia Weber's criticism of Robert Kohane's (.9@9) gatekeeping discourse on feminism as president of the International Studies Association is a case in hand. Kohane's test, classifying approaches to gender in IR as more or less useful to his own research agenda (framed as universal) was sharply criticised for the implication of perpetuating a discipline dominated by masculine discourses (C. Weber, .99G). On the Contextual:Internal side, a similar kind of criticism applies, only with a wider focus. Ashley's (.9@=) worry that neorealism crystallises the *status quo* in world politics is an example of theory-to-context analysis, whereas Curran's (5//9) study of the way political agendas may indirectly influence discourses of science in IR illustrates a context-to-theory approach. Both cases mobilise aspects of so-called 'critical theory' in a metatheoretical fashion, with a view to better understand the social and political roles played by theoretical knowledge in IR.

Our final way of metatheorising highlighted here is *historical*. It looks at the formation of IR theory from several perspectives. The Internal:Intellectual variety tends to follow a 'growth of knowledge' approach. The Great Debates narrative may be seen as a systematic account of the cognitive development and progress of IR as a discipline and exemplifies this historical way of metatheorising (Einhart, .9>Ga). Intellectual:Internal stories emphasise this 'growth of knowledge' with reference to how IR incorporated knowledge from other disciplines. The metatheoretical framework itself may be drawn from other disciplines too, as we see in the case of those using (peculiar readings of) Thomas Kuhn's notion of 'paradigm' in the field (Einhart, .9>Gb). On the Internal:Contextual side, one may study the impact of IR theory on the formation of the discipline, but most likely the account will move from disciplinary context to theoretical material. Tuomi's (5///) story of how constructivism gained acceptance into the field

by presenting itself as a *via media* between 'positivism' and 'relativism' is worth mentioning. #Sternal:6onte\$tual metatheorising in disciplinary history emphasises factors outside the discipline and their impact on theory. It can be illustrated by Arlene &ic!ner's (5//@) suggestion of a connection between economists of the *nited Bations #conomic 6ommission for Eatin America and the 6aribbean, political scientists reacting to geopolitical thin!ing under the influence of military rule and IR theory in Eatin America, especially during the 6old 7ar.

&hese five main ways of metatheorising 0 hermeneutical, corrective, evaluative, critical and historical 0 are not 'pure' types. &hey often overlap. ,or e\$ample, Cunne's study of the 1ritish 6ommittee and the #nglish chool plays both an interpretive role and a historical one, in that it helps us better understand the way 1ritish IR has developed under the influence of the chool. 7hile in many cases it will be hard to pin' point a single most salient element in a given piece of metatheoretical research, the fact still remains that these ways of metatheorising abound in the literature. IR scholars recognise these roles and associate them to metatheory, but many ta!e the additional step of denouncing metatheoretical research for one reason or another. 6ome, still, defend the way in which this discursive layer shapes the discipline. In this chapter, however, we are primarily interested in e\$ploring the implications of our basic definition of 'metatheory', as well as unpac!ing some of its general roles. Additional and specific e\$amples from IR authors, statements in favour and against metatheorising and further e\$plorations of how metatheory wor!s in our discipline in a selection of cases are left to the forthcoming chapters, in the second part of the thesis.

Synthesis and clarifications

#ach combination of 'focal point' (sub9ect'matter) and 'way' of metatheorising constrains and enables certain 'bridging' roles, indicated by the relations represented in &able 5.5 (above) and illustrated in the previous section. 6ome ways of metatheorising treat these specific relations with more facility than others, hence the importance of inferring the roles played by the possible combinations. After brea!ing down the topic almost to e\$haustion, in &able 5.F (below) I summarise what has been discussed so far.

Table 1: A 5x10 matrix of 'relations' and 'ways of metatheorising'

Focus	Relation	Interpret	#value	Refine	Critique	History
Internal:	$X \& h Y$	\$	\$	\$		\$
Intellectual	$Th_{IR} \supset Th_{IR}$	\$	\$	\$		\$
#Sternal:	$Th_{IR} = Ph$	\$	\$	\$	\$	\$
Intellectual	$Th_{IR} = Th_{<}$	\$	\$	\$	\$	\$
Internal:	$Th \supset B \supset I_{IR}$				\$	\$
Ontological	$I_{IR} \supset B \supset Th$	\$			\$	\$
#Sternal:	$Th \supset B \supset I$				\$	
Ontological	$I \supset C \supset Th$	\$		\$	\$	\$

Source: own elaboration.

Let me emphasise once again that the 'relations' addressed in metatheoretical arguments and listed on the table are concrete indications of metatheory's manifold bridging roles. They also account for its stress on 'theory', regardless of what else constitutes these relations. Theory (*Th*) is common to all such endeavours also regardless of their multiple roles of interpretation, evaluation, refinement, critique or simply historical account. For all their diversity, metatheoretical arguments still count basically as 'systematic discourse about theory'. After analysing the topic and its implications in a meticulous way, we can synthesise the findings once more and reaffirm this basic definition of metatheory, which does justice to the usage of the term in IR and cognate disciplines.

Ritzer's typology applies the Internal:#Sternal and Intellectual:Ontological continua to only those arguments which fit one specific kind of metatheorising. These go under metatheorising as a means for attaining a deeper understanding of (...) theory. In his view, they would not apply to metatheorising as a prelude to theory development or as a source of perspectives that synthesise theoretical material and overarch it (Ritzer, 1997, p. 6). In my analysis I have adopted a more encompassing role for those two continual axes and subordinated everything to the basic definition of metatheorising. The other varieties (metatheory leading to theory development and synthesising theory) happen to fit (albeit in different ways) into the combinations of both axes in the case of IR anyway. Moreover, I have also refined Ritzer's views in terms of the basic 'relations' between theory (*Th*) and the rest, highlighting the 'bridging' role of metatheory. This feature is often implicit in Ritzer's approach, but not as a general feature of metatheoretical research. Finally, I have also improved upon the original framework in my analysis by incorporating 'ways of metatheorising' into the

combination of several factors. Dreading down a subject to this level of detail is not, in my view, something to be pursued in every treatment of the topic, but the exercise is especially fruitful and original in this study, where the issue is being addressed as relevant in itself.

It is now time to clarify a very important implication to which some examples provided here allude: metatheorising can also be a self-referential exercise. (Metatheory, being a kind of theory, is potentially an object of itself. That is, because metatheory is a 'theory of theory', and because it is a kind of theory, it follows that it can eventually be employed in the (meta)theoretical study of metatheories. Thus, in addition to its more 'conventional' functions, it may well play the role of theorising about *metatheory* (Eaton, 1990, pp. 15ff). If the present thesis consisted in systematic discourse about *theory alone* (as in the case of part three), it would be metatheoretical. Despite operating mainly as systematic discourse about *metatheory*, this thesis is still metatheoretical.

Such an important point will be stressed again in my assessment of the strong critiques of metatheoretical research in IR, analysed in the forthcoming chapters (part two). There is another relevant implication of the basic definition to clarify. In principle, it would also make sense to apply the key 'ways of social theorising' seen in the previous chapter to another level, as 'ways of metatheorising'. The reason is very straightforward. Theories are artefacts. In this sense, they are part of the social world (Tafel, 1990). Patricia Jackson's analysis of the Vocational orientations of theory is a good example. She studies the different uses for 'theory' in IR depending on the 'ideal type' of persons employing theoretical perspectives and concludes that we gain a better understanding of what 'theory' means in IR by looking at whether they are interested in 'practice' or 'contemplation' (Jackson, 1990). Concrete metatheorising such as this case can be typical illustrations of how social theory is employed to analyse theory-in-society. It is, therefore, not a surprise that the 'ways of metatheorising' studied here are conceivably analogous to the types of social theory surveyed in the previous chapter. In sum, two points made explicit at this stage are: metatheory can study metatheory, and social theory in some cases can study theory.

There is, still, one further clarification to be made, pertaining to metatheory's position as a 'bridge' and its relation to philosophy and theory. The general distinction between metatheory and theory, as well as specific areas of overlap, are topics we have addressed already. However, what else can we say about the relation between philosophy and metatheory to further distinguish between them? (Metatheory would

seem to be a mere application of philosophy to a discipline. Philosophy is often portrayed as the 'discipline of disciplines', in that it cuts across special fields, always bridging between their specific and wider issues (Strauss, 1999). Can metatheory, then, be fully reduced to philosophy? The answer is not simple. Philosophers themselves tell us that philosophy is 'a way of thinking' (Warburton, 1996, p.), or 'a method of thinking, rather than a collection of facts' (Erigg, 1995, p...). Its questions are always of the broader kind: the nature of reality, the nature of knowledge, the nature of morality and so on. But similar issues appear in a narrower version in specific scholarly disciplines: the nature of social reality, of knowledge in social science, of the morality of social policy, etc. (Strauss, 1999, p. 59). Is the difference merely a matter of scope of the issues addressed, or does the bridging character of the application make a qualitative difference? Is metatheory only a 'method of thinking' or can it deal with theories as 'collections of facts' in a manner deemed as not so relevant to philosophy, but very relevant to 'science studies' or 'meta-science' (Fleiss, 1999; Radnitzky, 1997)?

According to what we have seen about the types and ways of metatheorising, we have to answer that, while there is some overlap between metatheoretical and philosophical inquiry, there are reasons to analytically draw a distinction between them. The simple reason is that metatheory often 'empirically' studies theory. This, according to what we have gathered above from philosophical writings, is very distinct from the task of philosophy. A more complex explanation is that qualitative differences may (depending on the case) be triggered by the 'bridging' operation. Scholars often imply that metatheoretical research is a kind of application of philosophical issues to a narrower field (e.g. Hollis, 1995, p. 15). There is a sense in which this 'bridge' simply adapts the issues to a special discipline (Lunge, 1999, pp. 1-2), but sometimes it operates as a productive mechanism. The bridge always involves 'theory' in the specific discipline, regardless of additional 'departures' and 'arrivals'. There is a wide perception that metatheorising constrains and enables certain kinds of theoretical argument. One often reads, for example, that 'diversity of forms' of IR theorising 'can be explained by different meta-theoretical commitments'. Such commitments 'can have a major impact on substantive theoretical traditions, currents of thought and on specific theories effectively 'shaping' them into new forms' (Urgensen, 1997, pp. 4-5). In this sense, say, Ullrich and Wright (1997, p. 4), 'meta-theoretical positions direct, in a fundamental way, the manner in which people theorise'. This kind of mechanism, argues Ullrich, is not something that is merely on the surface of an approach. It is ingrained (1997,

p. 10). (meta)theoretical commitments derived from philosophical perspectives limit what sorts of explanations are logically possible, but they do not operate in a deterministic way (Kenny, 1975, p. 10). The connection may open up (or close down) avenues for substantive theory and thereby exercise an important regulatory influence (Endt, 1999, p. 10). (meta)theoretical 'bridges' always point to theory in a specific discipline, and in many occasions trigger this productive mechanism, creating something new, and not merely a simple application of philosophy to a more restricted field. Thus, while some may even define 'philosophical theory' as 'theory about theories' (to, 1975, p. 10), meta-theoretical research does specific things for IR which philosophy by itself cannot provide.

Final remarks

This chapter concludes the first part of the thesis. It has identified, analysed and clarified a certain 'discursive layer' of scholarship that pertains to each discipline. Followed by illustrations taken from some of the literature in IR and parallel fields in social science, the definition of meta-theoretical research as 'systematic discourse about theory' has been expanded in its manifold aspects and implications. With reference to focal point, or subject-matter, meta-theorising always relates to 'theory', but goes beyond this intellectual element, also looking at contextual factors. With respect to the 'bridging' role of this discursive layer, a number of combinations has been detected. In our discipline, meta-theoretical research may look at IR theory in isolation or in relation to other IR theories. In addition to this, it can relate IR theory to philosophy or theory in other disciplines. Moreover, contextual meta-theorising can shed light on theoretical elements of IR from the perspective of social dynamics in the discipline or outside the discipline. Taken together, these possible 'bridges' and 'focal points' are not the only ways to characterise meta-theoretical research in IR. Using examples for each case, this chapter has also looked at 'ways of meta-theorising', which play the roles of interpreting, evaluating, refining, criticising and telling the story of IR theory. The analysis concludes with a brief clarification of the distinctions between philosophy and meta-theory. While the nature of meta-theoretical and philosophical inquiry overlap to a certain extent, we do well in avoiding a full reduction of meta-theory to philosophy. (meta)theory always points to theory within a specific discipline, and depending on the case it can be a mechanism influencing the way we theorise in that discipline.

The general question of the role of meta-theory in social science and IR has now

been addressed. The role depends on the combination of the types, 'bridges' and 'ways' of metatheorising. This, however, is only part of the issue treated in this thesis. We are also interested in the concrete ways in which metatheory makes a difference in IR. The next couple of chapters, in part two, deal with what IR scholars themselves claim to be the roles played by metatheoretical research in the field. Whether negative or positive, these views help us better understand the disciplinary dynamics of metatheory in IR. It is now clear that metatheoretical scholarship provides us a great number of intellectual services. Nevertheless, before deciding in favour or against the use of metatheory in IR, we should carefully weigh the relative merits of both types of claims as they manifest themselves in the discipline. Although I provide my own position already in the next couple of chapters, there is yet another step to be followed. We should not rest content with the mere assertion, made in the literature, that metatheory can affect the discipline by operating as a 'mechanism'. This crucial point should be further developed by analysing how metatheory operates in concrete IR studies and tests on different topics. In the final chapters, or part three, this will be pursued with the side effect of providing even more illustrations of the roles of metatheory. Such 'spillover' will be the case for the simple reason that the exercise involves 'systematic discourse about IR theory and metatheory'. The next step, however, is to look at the negative and positive views of metatheoretical scholarship in the discipline of IR.

PART TWO

Chapter 4

Negative Views of Metatheory in International Relations

Introduction

Having defined what metatheory as such is and what its intrinsic roles are, I proceed now to a critical evaluation of the IR literature on the roles of metatheoretical research. In the previous chapters I have explored the notion of metatheory as systematic discourse about theory. As a partial conclusion, we are left with a list of roles that different kinds of metatheory may play according to their inner features. However, we cannot completely reduce an investigation of the roles of metatheory in a given field to mere philosophical analysis. For one thing, metatheory may also play accidental roles. For another, these functions – be they accidental or not – do not emerge solely from the intellectual or content-related aspects of metatheory, but also from the fact that this 'discursive layer' is employed by scholars and groups of scholars with certain purposes. Metatheories should, therefore, also study contextual roles performed by theories. By the same token, my (metatheoretical) study of IR claims will only be comprehensive if their own contextual features are also given due consideration. There is, therefore, the need to go beyond what we can know via philosophical and conceptual inquiry and closely examine metatheory as it is treated in concrete instances in the IR literature.

The following chapters bring the discussion, initiated at a general level, closer to our academic discipline. By thoroughly surveying the relevant material, they purport to elucidate the field's overall perspective regarding what metatheory is and does. In the present chapter I examine several negative views of metatheory, taken from a survey of the literature on the subject. In the following chapter, a similar procedure is adopted, but with the contrasting focus on what IR scholars see as positive aspects of metatheorising and their constructive contributions to the discipline, both potential and actual. Although I primarily emphasise what has been, and is being, said about metatheory in IR, I also go to the vicinities of the discipline (e.g. politics and the social sciences in general)

asking whether what happens there is analogous to the debates in IR. &his I do for a simple reason. Both IR objections to metatheory and supporting views in its defence often draw on extra-disciplinary debates in order to establish their claims. &his means that, in order to better understand the discussion of metatheory in IR one should also look at the context in which the arguments are developed. IR comments on the subject are often made in passing, and shedding light on them with reference to their origins may help clarify the issues at stake.

In this chapter, a critical review of the key objections to metatheorising in IR is provided as a way of assessing perceptions of negative roles that metatheory plays or could potentially play. Although they may be grouped in different ways, I find the following headings informative enough: (1) intrinsic features of metatheory; (2) the complexity of metatheory; (3) the teaching of metatheory; and (4) the politics of metatheory. As I go through each of these points in detail, I assess their respective merits as portraits of the negative roles metatheorising plays in IR. I leave a stronger critique of the objections to metatheory in IR for a later stage, but at the end of this chapter I discuss serious internal problems with these views. &here is something to the negative view of metatheory, but ultimately it fails to provide any strong claim for the elimination of metatheory from the field. As a consequence, we also need to consider potential and actual constructive roles that metatheory may perform, a task which is left to the next chapter. For the time being, let me consider the key complaints about metatheory in IR.

Intrinsic features of metatheory

One of the strongest kinds of objection to metatheory in IR comes from the notion that metatheoretical discourse is damaging or at least irrelevant, given some of its intrinsic features. &his approach is inherently analytical in the sense that it relies on strict conceptual distinctions connected to disciplinary borders. For some time, social scientists have been discussing what falls within the scope of their specialised fields and, by implication, what kinds of debates are better left to other disciplines (e.g. Dohweder, 1990; Ziddens, 1999; Trauss, 1995). &he analytical objection to metatheory emerges from such boundary disputes as an attempt to KpurifyP IR from what it must *not* be doing. &he same kinds of claims have already been developed in politics and expanded to general social science. For this particular reason, it is useful to take this background into consideration as well. &here are, accordingly, two main types of

analytically-oriented formulations about why metatheorising in IR is a bad idea. One of them focuses on Intellectual issues about what the contents of KIR'tal!P should be. I will look in more depth at this view because of its current relevance as pivotal in a heated recent metatheoretical IR Journal debate ((onteiro D Ruby, 5//9a)^{5A} & the other kind of analytical claim emphasises the 6onte\$tual crystallisation of scholarly practices that revolve around whatever is considered to be KIR'tal!P. & the common argument of both Intellectual and 6onte\$tual sides of the e"uation is that metatheory is intrinsically distinct from what the content of disciplinary discourse in IR should be. That emerges from this negative view of metatheory is the notion that what passes as KIR'tal!P must change, and so must the institutionalised practices of scholarship. & this is a strong objection to metatheory because it represents an attempt to eliminate metatheoretical discourse from the field.

& the first analytical claim against metatheory is that it must be differentiated from theory, with several Intellectual implications of a normative character. & this idea has been elaborated to e\$haustion in John Dunnell's wor!s, starting with his complaint against what he called the Kphilosophisation of political scienceP (.9>9) and moving to an e\$pan\$ion of the same logic to political philosophy (.9@=) and, then, to social science in general. As he deals with the latter, Dunnell (.99@) postulates a series of sharply distinguished kinds of practices. A first'order practice (e.g. playing football) is intrinsically different from a second'order practice (e.g. writing about football). econd' order discourses embedded in second'order practices (e.g. what sports Journalists say about a football match) do not have any authority or direct impact upon first'order discourses and practices. In the case of politics, political debate may be classified as a first'order kind of discourse, whereas a model of political analysis pertains to a second' order category. ecause it has a second'order practice (theory'ma!ing) as its focus, metatheory would "ualify as an even higher order of discourse. Bow, if it is indeed the case that a higher order of discourse does not have any authority or direct impact upon a lower order (as in the case of Journalistic commentary on football compared to football matches), then, by inference, it follows that metatheory also has no authority or direct impact over the lower orders (theory'ma!ing and politics'ma!ing).⁵⁼ & the increasing subsuming of social science to higher orders of discourse has been inflating special disciplines with philosophical debates that, at the end, have no positive role to play in

^{5A} Others in the debate are Iohman (5//9), Chernoff (5//9), Jackson (5//9), <ur!i (5//9), Mercado (5//9).

⁵⁼ A refutation of the basic premise follows below, at the end of the chapter.

these lower orders. IR, by implication, should discuss Kglobal political practiceP while, instead, it has been dealing more and more with Kpolitical theory'ma!ingP (2unnell, 5/..).

According to this view, then, what are the problems of discursive'order conflation faced by IR8 &o begin with, there is a false illusion that metatheory (or philosophy for that matter) is capable of settling problems that are the ob9ect of lower' order practices. In short% in this view metatheory is not, and cannot be, theory. ,or metatheory spea!s about theories, not society. &herefore, those who thin! that metatheory helps shed light on social issues are simply mista!en. As Cavid Armstrong (.99A, p.FA>) says, metatheory Lis inherently parasitic. It purports to be able to ma!e Kreasoned 9udgementsP between rival paradigms but it cannot produce paradigms of its ownM.^{5>} In other words, Lit would be a great mista!e to believe that metatheory can provide authoritative answers about how to study international politicsM (chmidt, .99>, p....). +nly the theoretical level may spea! of first'order practices. And besides, even in this case (according to 2unnellPs logic) a *theory* of IR would have no dealing with the *practice* of international political action.

&his leads to the second problem% the "uest for higher'order foundations in 9udging, or in providing directives for, lower'order discourses is doomed to failure (Ea!e, 5/..). &his is said to be the case not primarily as a matter of coincidence (despite the fact that philosophy of science, for e\$ample, has indeed been ignored by most natural scientists), but rather as a matter of necessity, for one discursive order intrinsically does not interfere with another. &herefore, why should IR scholars give so much credit to higher'order specialists li!e 3opper, Ea!atos, <uhn, 1has!ar and others8 (oreover, why should they pay more attention to them if not even natural scientists (whose practices their higher'order discourses originally emphasised) do (cf. 7ight, 5//5, p.5A)8 And, on the other hand, why should IR theorists be guided by philosophers of science and metatheorists if, in turn, theorists themselves are often ignored by politicians and other practitioners8 K&ownP, it is argued, does not re"uire KgownP in order to succeed. As sung in *Iolanthe*%

Then Britain really ruled the waves
(In good Nueen lessPs time) 0
&he -ouse of 3eers made no pretense
&o intellectual eminence,
+r scholarship sublime;

^{5>} &his claim is both logically and factually wrong. ee my discussion below.

Yet Britain won her proudest bays
In good Queen Elizabeth's glorious days⁵⁴

Discursive order conflation is actually a concern not only for Zunnell, who certainly poses such questions at both necessary and contingent levels, but also for those who partially agree with him. Schmidt (1997, p. 10), for example, alludes to Zunnell's (1990, pp. 10-11) complaints about metatheorists' pretensions to judge theories, but concedes that metatheory might contribute to the development of a good theory.⁵⁹ The main problem according to him, however, is that many metatheoretical claims, rather than simply offering commentary on a given theory or set of theories, seek to provide the foundations for building theory and for making authoritative judgments about appropriate knowledge in political inquiry.⁶⁰

From the orders of discourse perspective, then, IR can never be metatheory and metatheory can never be IR. The idea is that we have been conflating distinct orders of discourse and that this has generated a great deal of unnecessary friction and, worse still, unnecessary scholarship that already starts from the wrong premise that discursive orders can indeed cross-fertilise (see Collins, 1990). We must step back from this procedure and purify KIR's tale. For that, what we need is a sharper distinction between metatheory and theory: the clearer, the better.

Although Zunnell (1990) himself only marginally touches the subject when he discusses metatheoretical IR, (Monteiro and Ruby have applied a similar reasoning to the field of IR with reference to the same approach. In their protest against the use of philosophy of science in the field, they describe most IR metatheoretical debates as "quests for a foundation to ground the discipline. Such foundation would be able to address issues like validation of research, basis for graduate training following standards of good scientific practice, criteria for faculty selection and so on. According to the authors, the intention of those engaging in this kind of conversation is to settle the science debate once and for all in the pursuit of a single foundation to define IR's relationship to science" (Monteiro & Ruby, 1999, pp. 1-2). Such aspiration is seen as an imperial project in that certain philosophical worldviews, when applied to IR in a

⁵⁴ An operetta by Sir Arthur Sullivan. Eibretto by T. S. Eliot (1951). See below for more on politics and policy.

⁵⁹ Schmidt (1997) displays not only his indebtedness to Zunnell's work and influence, but also (ironically) great skill in metatheorising in a historical and hermeneutical fashion.

⁶⁰ If we take together both statements, we can see that Schmidt does not manage to reconcile metatheory's eventual ability to develop good theory on the one hand and, on the other, its intrinsic incapacity to ground theory-building. Besides, as shown in the previous chapter, this is not the only role of metatheory.

metatheoretical way, attempt to KcoloniseP rival approaches with their own view of reality and !nowledge. imilar claims about this KfoundationalistP intention behind metatheoretical debates in the field are repeated at several points. IR scholars are depicted as emulators of philosophyPs discussion on what would be Lan unsha!able foundationM (p..9) committing themselves to certain philosophical stances by portraying them Las *the proper foundation for IR as a /hole*M (p.55, original emphasis).^{F.} (oreover, despite the fact that metatheoretical debates were Lsupposed to solveM the foundational issue (p.5G), they still Lcontinue their "uest for an unsha!able philosophical foundation, capable of settling once and for all the most fundamental "uestions in the field 0 its scope, goals, criteria, standards, and methodsM (p.55).^{F5} In this account, then, metatheory is portrayed as playing in IR the negative role of attempting to bring uniformity to the discipline by trying to subsume it to certain philosophical KfoundationsP.

-as IR managed to achieve a final and irrevocable decision concerning metatheoretical issues8 -as this !ind of debate reached the Kimperially foundationalP goal attributed to it by (onteiro and Ruby8 A negative answer is the neSt step in their argumentation. &hey proceed to a highly critical diagnosis of the attempts to connect philosophy of science to IR. In the first place, metatheoretical discussions in IR Lhave contributed to the fragmentation of the discipline along meta'theoretical linesM (p..>). &his is seen as an intrinsic effect of strict commitments to specific philosophical formulations%

&he multiplicity of reasonable foundational positions, supported by prominent philosophers of science, virtually guarantees that each contending perspective in IRPs KscienceP debate will find a well'supported philosophical position to 9ustify its particular approach (O). 6ommitments to Instrumentalism, ocial 6onstructivism, and cientific Realism now anchor a three'cornered fight between positivists, anti'positivists, and post'positivists, with each side wielding its favored philosophical position (O). It should therefore come as no surprise that the turn to philosophical authority in IR has failed to produce a consensus on the proper foundations of the field ((onteiro and Ruby 5//9, pp..='>).

F. &he authors echo 2unnellPs view of the intentions of philosophy of science as related to lower'orders scientific practice.

F5 &hese claims about the Kimperial pro9ectP of specific philosophies involved in the metatheoretical debates are never demonstrated. Bor do the authors refer to any further literature or any other source (e.g. metatheorists in IR claiming they want to establish a single foundation). Co these affirmations, then, emerge from sub9ective 9udgement based on the authorsP personal e\$perience in the discipline8 &he disputable character of the claims is rhetorically mas!ed in a plethora of repeated statements about how metatheoretical debates in IR imply Kimperial pro9ectsP (ee (onteiro and Ruby 5//9, pp.55'5F; GF besides the passages "uoted above).

The irony is that, according to them, the whole point of metatheoretical engagement is to bring a definite end to such fragmentation in the discipline instead of making the situation worse. For this reason, it might be claimed, as Friedrich Schlegel (1801, p.5) puts it, that this project of securing knowledge through (O) finding absolute foundations failed. In sum because there is no agreement between the several schools of philosophy, looking metatheoretically for philosophical authority to rule over IR can only lead to sharper antitheses in the field. This is what has been called the inherent foundationalist tendency and problematic nature of metatheoretical arguments (Schmidt, 1997, p....).

In addition to the fragmentation brought about by the failure of the so-called Imperial Project, another negative role played by metatheory in IR according to Monteiro and Ruby is the rejection or endorsement of specific IR contributions due to their implicit or explicit philosophical background.^{FF} That is to say, regardless of an approach's substantive (i.e., lower-order) claims about the practice of world politics, agreement and disagreement from other researchers emerges primarily in terms of metatheoretical issues.^{FG} There has been, then, a shift in the criteria of disciplinary appraisal of scholarship. From judging work on its substantive contribution, IR has increasingly been making *a priori* judgments based on foundational commitments on what constitutes legitimate work in IR and based on the degree to which the chosen approach conforms to a particular conception of science (Monteiro & Ruby, 2009, p.10). Defending a similar opinion, Schlegel (1801, p.5) portrays the joint effects of disciplinary fragmentation and (mis)judgement due to metatheoretical clashes as leading to a virtual breakdown of academic dialogue. Metatheory implies not only growing disagreement, but also 'misunderstanding' or lack of hermeneutical charity.

Cisagreement is not the same as misunderstanding. The latter is indicated when people speak past each other. When is this the case? Usually, it is when people believe they are right and all others are wrong. It is seen in the habit or stance of arguing that my methodology or, worse still, my metatheoretical preference is valid, whereas yours is not.

Due to its inherent tendency to intrude in alternative orders of discourse, therefore, metatheory is blamed for bringing to IR scholarship much fragmentation and less

^{FF} In this respect their argument is an extension of Zinnell's view of what happens with metatheory in social science as a whole. See above.

^{FG} Schlegel's (1801) classical dismissal of Reflectivism based on the latter's failure to conform to the mainstream view of scholarship comes to mind.

incentive to cooperate in cross'approach endeavours. Instead of evaluating propositions according to their capacity to shed light on international political phenomena, IR scholars have been (mis)driven by metatheory to judge themselves for their ability to explain, affirm or deny certain features of scientific practice. The only solution, again, would be to provide a clear'cut distinction between discursive orders, thus eliminating metatheory's pretension to invade the realm of theory.

If the Intellectual objection to metatheory complains about its negative roles in conflating KIR'tal!P with Ktheory'tal!P, the 6onte\$tual claim, on the other hand, is a statement that metatheory, being different from KIR'tal!P, should be relocated as an institutionalised practice. This second kind of analytic attempt to eliminate metatheory from the field also draws a thick line between metatheoretical and substantive research. However, it takes the further step of drawing another line between both as they are actualised also in practice. If Ktheory'tal!P is not the kind of activity that IR scholars normally (should) do, then, in the current university curriculum, it must also be taken away from IR, even if it may be valuable elsewhere. Halliday (.99A p. >GA), for one, illustrates this position. He acknowledges, to some extent, the pertinence of metatheoretical aspects of scholarship% Lmethodology and KmetatheoryP are important for social sciences, IR includedM. However, he has a different view about the ideal institutional setting of such discussions% Lthey should be discussed where they belong, in philosophy departmentsM. Best we miss the positive properties of metatheory, he also recommends Lwriters on IR to be more aware of, and (...) students to be more literate in, the philosophy of the social sciences in generalM. Just how it would be possible to isolate metatheory from IR both Intellectually and 6onte\$tually and still require IR scholars to get acquainted with metatheoretical material is a puzzle that Halliday leaves us to grapple with.

Chris Brown (5/.5), Hincant Bouliot (5/./, p...'9.) and others (Kratochwil, 5///; Bouliot D Adler, 5/..) further develop the 6onte\$tual side of the distinction between discursive orders. They all emphasise the contrast between the practice of diplomacy, statecraft and other aspects of world politics, on the one hand, and the practice of theorising IR on the other. They stress the differences between intuitive and 'background' ideas orienting the practice of world politics and those guiding IR scholarship. A clear dichotomy, therefore, is postulated between 'IR'tal!' and 'IR' practice'. Brown (5/./, p.=@F) tries to further develop the 6onte\$tual claim by extending it to the difference between theory and metatheory. He highlights the

asymmetry of 'academic expertise' between philosophers and IR scholars. Institutional isolation, reflecting an ideal discursive distinction between metatheory and KIR'tal!P, is even more attractive in a context where 'abstract research' in IR has turned into a very demanding kind of specialisation and perhaps should be left only to those who have mastered a complex and difficult body of literature which non-theorists may well find daunting.^{FA} The claim behind this intention to drag metatheory out of the discipline is that metatheory has been playing the negative role of distracting IR scholars with philosophical quarrels that, interesting as they may be, should be discussed elsewhere. And this must be the case because metatheory is, intrinsically, *not* IR, and vice versa. This summarises the contextual claim, which is actually just the institutional application of the Intellectual or analytical distinction between KIR'tal!P and Ktheory'tal!P.

The complexity of metatheory

In addition to the negative role played by metatheory due to its intrinsic property of conflating discursive orders, there is also a complaint against metatheory due to negative KexternalitiesP that emerged in IR (supposedly) out of an increase in metatheoretical activity. In this section I deal with problematic effects of metatheory related to the field's fragmentation, to the 'questionable' quality of metatheoretical research and to practicalities of metatheory. I have grouped them together under this heading because the notion that metatheory is KcomplexP and has led to further KcomplicationP of IR is typical of the positions addressed in this section. A discussion of the remaining externalities is left to the rest of the chapter, where we see the perception that an increase in metatheoretical activity has negatively affected teaching as well as some political issues connected to the role of the discipline in society.

The first way in which metatheory is said to KcomplicateP IR bears some relation with (onteiro and Ruby's (5//9, p.6F) final verdict on the failure of philosophical debates in achieving a consensus on the ultimate foundations of the field% Foundations have not provided any firm bases for standards, have not settled any methodological debates XandY have not conclusively decided any empirical 'questionM. & to the contrary, metatheory has actually made things even worse by reifying the unproductive division of the field along artificial fault'linesM, leading to lac! of dialogue and constructive

FA See below for more arguments about negative outcomes of metatheorising in terms of practical aspects of scholarship.

criticism. Using a geographic analogy, Rose (1993, p. 11) achieves a similar conclusion. The discipline of International Relations has developed into an ocean in which distal archipelagos, themselves subdivided into islets of theory, provide abode to aborigines who speak with each other in their local tongue, using mutually reassuring codes. Holsti, to whom I have referred before in connection to the fragmentation of IR, portrays the phenomenon as an outcome of the clash of different metatheoretical orientations since the times of the earlier debates.

There are always debates in the field. There were immense debates in the 1970s between the more traditionally oriented scholars and the so-called behaviourally oriented scholars. These debates were sometimes vicious; people really abused each other, even on a personal level. This kind of cold warring infected departments and critically damaged the trust and intellectual pluralism that are the hallmarks of academy. The postmodernist challenge has led to the same kind of behaviour; the rejection of antecedent knowledge, the attempt to create something new, to question all that has gone on before, and to suggest that it was all wrong.

Abuses from one side against the other and the threat to intellectual pluralism, according to him, should be enough to counterbalance what some may see as a kind of excitement coming from metatheoretical clashes. You might say that intellectual warfare is dramatic but I do not think it necessarily leads to a very happy or productive intellectual community. Instead of bringing more agreement to IR, therefore, metatheory has made the discipline even more complex and fragmented (interview with Rose, 1995, pp. 55f).

However, metatheory is not alone in performing this role of augmenting intra-disciplinary fragmentation and insulation. In the past, a similar accusation was raised against any theoretical or philosophical activity as such. The insulation of theorists from non-theorists in fact led David Singer (1975), who saw himself as bridging between one extreme and the other in his scholarly practice, to complain about the two-culture problem in IR academe. That is, instead of portraying the field's internal insularity as a function of subscription to specific theoretical or metatheoretical approaches, Singer located the problem in the fact that many empirical researchers would simply ignore what was being produced in the theoretical realm, whereas theorists lacked the ability and will to look at historical and contemporary concrete issues. In short, the two-culture thesis, as an alternative way to look at IR's fragmentation, ascribed the latter's origin to the specialisation that emerged from the abstract character of theorising. On a

feature that has been extended to metatheory.^{F=} When Brown (5//, pp.=>9'=@G) recycles IngerPs notion, he alludes to metatheoretical activity in the current KtheoristsP group. Speaking about the British IR community, he laments that Ltheorists have come close to achieving a situation in which they have marginalised everyone elseM (p.=>@). &he perception is that the same applies beyond the confines of British IR. In a speech delivered to IR scholars, William Wallace (.99=, p.F/F) depicted the cult of abstraction as a general trend in the social sciences%

&here is a tendency for all academic disciplines to demonstrate their intellectual standing in the university by privileging theoretical studies over applied. &his tendency has been particularly evident in the social sciences, as they struggled to gain respect from disdainful classicists, philosophers and natural scientists.

With this division over Kabstract vs. empiricalP research, non'theorists, in turn, run the risk of being driven by what Brown (5//, p.=@A) calls 'a generalised anti-intellectualismM. &his, he argues, is 'an unhealthy situation from pretty well every perspective, not least that of the theorists themselvesM (p.=>@). It may be granted that metatheory alone did not give rise to the Ktwo'culture problemP, but it should be held at least partially accountable for it.

Fragmentation leading to lack of dialogue is not the only negative outcome emerging from metatheoretical activity. &here is a second category of perceived KessentialitiesP connected to the fact that metatheory is a particularly demanding specialisation, namely, the bad 'quality of metatheoretical scholarship in IR. &his can be noticed mainly in uncritical attempts by IR scholars to KimportP certain ideas developed outside the field and to adapt them to the new context. Speaking in terms of past and recent social science as a whole, Dunnell (.99@, p.F=) complains that theorists have insisted in subscribing to certain philosophical positions, yet Llittle or no considerationM was given 'of the extent to which these philosophical accounts were adequateM. Referring to the metatheoretical adaptation of KforeignP knowledge to the discipline of IR, (artin -ollis and Teve mith (.99., pp.F9F'G) alert us to Lbeware of gurusM%

It is tempting to import positions from elsewhere in the social sciences, complete with the sacred names of theorists associated with them. &hese gurus may be too simply credited with a coherent analysis which solves the problem or at least indicates where a solution will be found (O). &he danger is not that the importers believe in magic themselves but that they break into previously accepted theory by means of compressed survey articles which identify gaps tailor-made for the new guru. &he imported

F= A similar tension occurs between 'quantitative' and 'qualitative' researchers.

positions are then presented in a summary form which suggests that they have only to be stated to be believed.

Those who metatheoretically import knowledge from other specialised disciplines and from philosophy run the risk of decontextualising and oversimplifying the ideas (see Busset, 5//@), and so do the others who critically reflect on such moves%

It is a lot easier to make fun of figures such as Cerrida Xhermeneutics and languageY or Eacan XpsychoanalysisY (or Rawls Xlaw and ethicsY or chmitt Xpolitical theoryY for that matter) than it is to master their thought 0 it might be that having done so one would still wish to put the boot in, but it would be nice to think that some of the people who are readiest with their criticism had actually done the necessary leg'work!% all too often this is not the case (Irown 5//=, p.=@A).

Similar complaints against both types of protagonists 0 importers and critics 0 have been voiced in the specialised metatheoretical literature itself. Discussions at this level have often digressed to e\$egesis of e\$tra'disciplinary material (Cornprobst, 5//9, p../F), and one of the reasons presented is that IR scholars tend to be bad philosophers and natural scientists. Thus, whenever such trans'disciplinary trends are at play, there is always the risk that IR metatheorists will adopt highly contested views, as they already have in the case of <uhn, Ea!atos, Ihas!ar and, surprisingly, Carwin (Iell, (acConald, D &hayer, 5//.; &hayer, 5///, 5//G) and a few palaeontologists (6hernoff, 5//@; Han Ielle, 5//=)S In any case, prolonged discussions on marginal topics are clearly outside the scope of IR and, on the other hand, in their respective fields they would probably be too basic and unattractive. Having complained about the disciplinary fragmentation caused by the rising volume of metatheoretical discussions, Iu?an and Eittle (5//., p.F5) add that these debates Lhave been imported into IR from other disciplines, reproducing intellectual oppositions formed in the past and within different contexts, lacking very often the e\$pertise that gave rise to them in the first placeM. This means that they do not have much momentum per se and that, whenever they occur in IR, they may lead to more confusion than enlightenment (see -olden, 5//5, pp.5=A'>/ for e\$amples).

Critics of metatheory also point out an additional component that comes from these patterns of Kmovement of ideasP. They correctly identify a time'lag between the initial discussions and their repercussion in IR. -ollis and mith (.99., p.F9F) observe that Ltheoretical discussions have often come late to international relationsM. LIn what seems to be a recurrent patternM, says chmidt (.99>, p../>) more elo"uently,

Controversies that once held sway over other academic disciplines come belatedly to international relations. Michael Bicholson (1995, p. 9) has a similar impression:

With the time-lag which it seems that we require in International Relations before catching up with the intellectual fashions which convulse the rest of the chattering classes, we finally caught up with the problems of (O) philosophy in our area too. It seems that we are all eager readers of the *De/ For% Revie/ of Boo%es* but only get our copies ten years late.

This occurs, according to Luhan and Eittle (1995, p. 5), due to an inversion of priorities. IIR is placed in a peripheral zone, where issues that have been fought over and sometimes resolved elsewhere, arrive late and clash through a constituency that is not all that well trained to deal with them. Suggestively, part of their essay's title asks 'Why International Relations has failed as an intellectual project' (p. 9). In sum, by bringing extra-disciplinary ideas too late to the field, and due to the poor quality of their status, metatheoretical discussions play the negative role of shifting the discipline's efforts to marginal issues as if they were priority. As Brown (1995, p. 5) observes, it takes IR research some way away from its origins.

So far, I have considered under this heading problems of disciplinary fragmentation and problems related to the quality of metatheoretical scholarship in IR. There is a negative perception of metatheory attached to the fact that both kinds of Kernalities bring more complexity and confusion to the field. I now turn to the third and final category of complexity-related claims, which encompasses several practical issues of metatheory. One objection to metatheory based on pragmatic (and, sometimes, subjective) considerations such as lack of attractiveness and lack of clarity in metatheoretical discussions. These perceived properties of metatheory, therefore, play the role of hindering the development of potentially constructive interest in world politics by introducing unexciting and confusing notions into the discipline of IR.

First of all, there is the opinion that metatheory is one of the most boring issues that could possibly arise in social science. For decades, this judgement has been evolved at pivotal moments of the discipline. Mans (Orghenthau (1995, p. 7), for example, alludes with little excitement to the argument over ways of theorising which divided IR into a dualistic 'new scholasticism' comprising traditionalists (see Hull, 1995) on the one hand and behaviouralists (see Kaplan, 1995) on the other. In a review of the same discussion, Fred Beal and Bruce Jamlett (1999, p. 5) denote that they share the same opinion: 'Scholarship about scholarship, like theory about theory, and teaching about

teaching, is rarely very stimulating. They find it lamentable that such commodities are not in short supply in the field of international relations. That is, in their view there should be less meta'scholarship, metatheory and meta'teaching. Despite this, since that occasion the volume of metatheoretical contributions has increased and contemporary scholars have manifested a similar opinion regarding the unexciting character of metatheory. Brown (1997, p. 6A), for one, labels current discussions over science in IR as 'ultimately a series of not very interesting metatheoretical debates'. Having interviewed several theorists about ways of identifying themselves in the discipline, Porter (1997, p. 6F, my emphasis) reports%

Several respondents indicated that they either did not like or were not entirely clear on the significance of this "question. Rob Walker, for instance, added two other categories, 'theoretically interesting' questions and 'politically interesting' questions and ranked them well above the other categories. I.e., nationality, IR subfield, issue/area, major IR paradigm, key concept, regional focus and epistemology: methodology. Adding that 'it is not all that difficult to identify worthwhile scholarship whatever the approach. Far too much energy is spent classifying such approaches' usually in a very crude way' and far too much boring research and writing is produced as a consequence.

A similar opinion occurs in the broader scope of social science as a whole. What Epiped (.1999, p. 5F=) describes as 'a prescription for a rigorous philosophy/avoidance strategy'. Caly (1999, p. A>, my emphasis) summarises the issue% 'Engagement with ontological and epistemological issues in metatheoretical political study has been arguably *less than fully loaded*. Too often it seems that they are treated as *unpleasant hurdles* to be quickly vaulted in order to get on with the business of political analysis. (Metatheoretical debates seem to play the negative role of a nuisance in such situations and, more importantly, when they have to be taken seriously this is regarded by critics as an excruciating task).

Then it comes to practical issues, besides its lack of attractiveness, metatheory is also perceived to involve discourse of a significantly confusing quality. This applies primarily to the language often employed in metatheoretical debates, despite, of course, their inherently high level of complexity. Such 'eternality' is paradoxical, considering that one of the key arguments in favour of metatheoretical discourse is its potential to help us make sense of what goes on academically in IR. The vast theoretical outpouring in the field means that students beginning the discipline today are faced with a bewildering variety of theoretical perspectives (Armstrong 1999A, p. FA=). Given this context, metatheory has been presented in textbooks (Lurchill & Einlater, 1999A; R.

Donald Urezen, 1997; Donald Wright, 1997) as a helpful tool that deals with some fundamental problems of cross-theoretical debates (Smith, 1997, p. 5). Despite this, whenever abstract knowledge is implied, great care should be taken in order to make its output available to non-specialists in this kind of research (Brown 1998, p. 10). Amongst those blamed for being less than precise in their communication of ideas are critical theorists and poststructuralists, both of which draw heavily upon philosophical literature in their argumentation (Alliday 1999A, p. 19). However, the same charge may be made against mainstream works (e.g. Morton 1999A case for adapting systems theory to international politics).^F Regardless of its particular origin, complex language is better avoided. We ought to neither hide our knowledge in obscurely erudite terminology, nor to lose ourselves in scholastic word games (Wallace 1998, p. 10). As Colsti (1999, p. 59) says, In many of our scholars' excursions into epistemology, ontology and metatheory, we sometimes lose sight of one of our common purposes as international relations scholars% to make a seemingly difficult and often chaotic field (O) more intelligible.

Besides the paradox of being meant to simplify things and ending up complicating them with the use of obscure language, metatheory has also led to a further negative externality% its occasional alienation from those non-specialists who want somehow to have access to metatheoretical knowledge. As an example, notice how Stephen Han *et al.* (1997, p. 5) attempts to explain why he wrote a (partially) metatheoretical research guide without much reference to the metatheoretical literature%

I learned some important things from writings on philosophy of science and social science methods, but I have found much of that writing abstruse and unuseful. It was often easier to invent my own answers than dig them up from the realms of muddy arcana produced by philosophers and methodologists, even when the answers existed somewhere in those realms.

At a later point, he avoids further engagement with the metatheoretical literature in a disclaimer on why he does not wish to pursue a detailed exposition of a set of writings in the philosophy of science that are "quite relevant to his topic. The arguments, he claims, are well hidden in tortured prose that gives new meaning to the phrase badly written and no reading of such dreadful writing is ever certain or final (Han *et al.* 1997, p. 10 note 11). (Maybe Han *et al.*'s choice of ignoring the core literature as he

^F For an analysis of complicating rhetoric, see critiques by Pegele (1995) and, more informally, Walt in an interview (Alliday & Rosenberg, 1999, p. 10).

writes about metatheory contradicts accepted academic standards, but in any case one should not expect such a highly specialised vocabulary to be universally known to all non-specialists, even if they are part of the field. The cultivation of a complex language, of obscure terminology accessible only to those already deeply immersed in the specialist literature, is a justifiable tendency only in theoretical writing within specialist journals (Wallace 1998, p. 16/A). If metatheory plays the negative role of alienating itself from non-metatheorists due to the complexity of its issues and presentation, it poses an even greater problem for non-scholars. This particularly applies to the case of students, on the one hand, and, on the other, policymakers and additional practitioners. Our student audience – at least, our undergraduates – need something more straightforward; and a wider audience will remain beyond reach unless addressed in terms which they can understand without too much difficulty (Wallace 1998, p. 16/A). These two issues – metatheory in education and metatheory for the wider audience – are the specific themes of the next two headings.

The teaching of metatheory

Teaching is one of the central aspects of the communication of metatheoretical knowledge beyond the research community that must be addressed. Although the IR literature focusing solely on instruction which specialises on metatheory is underdeveloped, the topic has figured sporadically in journal articles. In one of them, Stefano Zucconi treats metatheory as paramount to the IR curriculum and deals with it predominantly in a favourable way. However, he also warns us that, if badly taught, metatheory could lead to certain negative consequences. For example, depending on the module's or programme's structure, it risks moving the discussion too early to a highly theoretical level and leaves a certain taste of *ad hoc*-ness to it (2001, p. 100). Moreover, he argues that courses focusing on meta-theoretical differences between approaches may have the disadvantage of being taught in a top-down manner. This is likely to lead to excessively passive learning. If, instead, instructors opt for a course design emphasising active and critical learning through several assessments, they must be ready to give very thorough and frequent feedback to students, which of course is a time-intensive strategy and may not work well in big settings (p. 101). In a more empirically-oriented article that compares different styles of teaching from the point of view of learners, Ariacovich (1998, p. 102) reports that even groups of students who were familiar with the scientific jargon and the different dimensions of

the discipline (O) were nevertheless puzzled and in some cases confused by the analytical abstractions. Referring to a module that ascribed great weight to debates over research methods and approaches (traditional and scientific methods, case studies and comparative methods, content analysis and event data, forecasting, and theory and policy making) (p. 9), students complained that they could not appreciate the entire structure of the theoretical edifice and that they felt lost in the (O) logical and theoretical abstractions of the lectures (p. 9). With reference to another module, which was driven by a strong methodological and epistemological concern with the scientific approach and which included issues like science and wisdom (p. 10), students regarded the course material as too abstract and detached from reality, and expected a more comprehensive and concrete discussion of international problems (p. 9).^{F@}

Speaking from a more critically oriented perspective in a symposium on IR teaching, (Cowan and Bel (1995, p. 5A), editors of a textbook that specifically targets those interested in Africa as a region, declared that students everywhere struggle to define contested concepts clearly, and to engage in theoretical discourse in which they discover and interrogate their own assumptions. Surprising to encourage reflection, the authors have adopted a student-centered, outcomes-based approach. In such method, the emphasis is not so much placed on what students should know, but rather on what they should be able to do once they have completed a module or course. It starts with what the student knows and tries to develop both a deeper understanding of that prior knowledge, and a broader application of this deeper understanding. In emphasising an intuitive understanding of their region as a base from which to explore the dynamics of the world as a whole, they have clearly tried to follow a critical pedagogical approach (see 3. Aire, 1995) and a non-Western view of IR. Commenting on their work (Bel and Cowan, 1999) and on the potential contribution of a similar teaching method to IR in Eastern Europe, Qatilo Zabi (1995, p. 5A) laments the tendency of some universities to accept Western expertise and textbooks without much reflection on whether or to what extent these works correspond to the actual needs of (non-Western) students. Donald Zordon (discussant in Loyer et al., 1995, p. 5F) also endorses a more critical look at IR pedagogy and its implied avoidance of the abstract debates taking place in the West. Welcoming Bel and

F@ 2000 and 2001 concerns about careful planning of their role as teachers in metatheoretical modules reflect, even if indirectly, a specific emphasis of certain streams of educational theory on the teacher as a mediator. See Redler and Shields (1999) for a recent overview of Eev Hygotsky's classical statement and Foucault, Cepresbiteris and Machado (1999) for an account of Feuerstein's contemporary and influential formulation of the topic.

(Cowan's textbook, he praises its Lshift away from looking at the world through the theoretical Keyes of the industrial West. & the extent that metatheoretical debates reflect an academic interest that does not correspond to the local intellectual needs, does not add anything to the local disciplinary context and, moreover, distances the students from more pressing needs related to their own political agenda, metatheory may be a hindrance to learning.

There are additional problems connected to the pedagogy of metatheory as seen from a critical perspective. Because the disciplinary discussion on this matter only touches some of the issues, it is perhaps more illuminating to bring to evidence a similar debate recently introduced in the broader literature on politics. Concerns related to disciplinary diversity and to the instrumental role of metatheory in the acquisition of critical skills are pivotal in a review essay by Stephen Bates and Laura Benins on negative educational outcomes of metatheory. In their article, they ask whether the contents of both political analysis textbooks and teaching introduce metatheoretical discussions in a way that provides students enough incentive to think independently. Bates and Benins (2009a, p. 11) identify metatheory as a relevant educational tool in the context of a disciplinarily diverse field of politics. According to them metatheory is an important element of political science, as it helps students to appraise, differentiate and choose between competing philosophies, theories and analytical traditions. Eileen Benins and unlike the commentators of Bates and Cowan, they appreciate the potential didactic benefits of highlighting central metatheoretical debates to make sense of theoretical and philosophical plurality. However, they are even more concerned than the African textbook authors with the ability of metatheory to bring about reflective learning. Drawing on Paulo Freire's (cited in Bates and Benins 2009a, p. 11) view of Kmentoring, intended to give rise to the possibility that the students become the owners of their own history, they indicate that metatheory should be instrumental to this emancipatory goal.

As such, teaching and learning within political science should not be about the instruction and regurgitation of knowledge. Instead, teaching should mean alerting students to different ways of thinking. It should provide a non-prescriptive basis from which students can reflexively engage with the material in order to uncover relationships, connections and underlying patterns and, consequently, participate in critical analysis (pp. 11-12).

Such engagement with metatheoretical issues is regarded as a way of Kmentoring

because students could develop a theoretically informed, innovative and research-oriented disposition (p. A=). In short, metatheory could play a positive role in educational terms.

The problem, though, is that the teaching of metatheory in politics has been failing on both didactic and emancipatory fronts. In the first place, as the previously discussed case of IR illustrates, metatheory may (and, actually did) bring more complexity and confusion to disciplinary contexts of theoretical plurality. After recognising the positive side of teaching metatheoretical issues in politics, Latour and Woolgar complain that the main available textbooks and predominant educational methods actually hinder intelligibility. According to them, the way metatheory has been taught is pedagogically problematic (1990b, p. 5/2), as the dominant literature on ontology and epistemology within political science retains some unacknowledged and problematic assumptions (1990a, p. AA). Apparently there is a trade-off between accessibility and precision. They agree with Rorty and with Macneil on metatheory-related lecturing being particularly demanding in terms of balancing both sides of the equation. Latour and Woolgar (p. A9) identify problematic issues in the most widely adopted texts. One particularly striking example is the frequent conceptual conflation of epistemology and ontology. If textbooks are inconsistent and inaccurate as they say, then the goal of clarifying differences and disputes between political scientists (which) are more profound than students often assume (p. A>) will hardly be reached. Consequently (...), the very activities promoted by an understanding and appreciation of ontology and epistemology, are paradoxically undermined and become stunted at this metatheoretical level (p. A=). One of the first negative roles played by metatheory in educational terms is that it may actually make things worse when, instead, it is supposed to clarify and elucidate.

The other negative outcome of teaching metatheory in the 'wrong' way is that, despite its alleged mission to encourage reflective learning, in the pretension of presenting one single argument that would define the disciplinary situation for the students, metatheory actually leads to closure. This, in turn, prevents the development of an adequate attitude towards the mentoring process. Closure is damaging to emancipatory education, according to Paulo Freire (cited in Latour and Woolgar 1990a, p. =.), because mentors are neither supposed to encourage their ideas nor provide closed answers to be reproduced in the mentees. Having analysed a few introductory texts to political analysis (e.g., Jørgensen 1990; Jørgensen, 1995), Latour and Woolgar

(5//>a, pp.A9'=/) "uestion the way in which they convene ideas about the relation between ontology and epistemology as metatheoretically applied to politics. uch te\$t\$, they argue, portray the relation as a Kcase closedP% either one precedes the other or both go together. -owever, these are not the only options available, as becomes evident when we observe neighbouring disciplines li!e sociology and IR (pp.='=.; see also Caly 5//@). 1ates and 4en!ins (5//>a, p.=.) believe that KclosureP Lmay lead to an outcome where students are more prone to support and defend particular theories, and re9ect, find contradictions within and fail to understand othersM. As a result, Ltheir meta'theoretical choices are hindered by the position of these teachersM. 2u??ini (5//., p../9) refers to a similar situation involving metatheory and IR. &here is often La confusion of ideologies with meta'theoriesM, which leads to Lthe ris! of tal'ing these clusters in purely ideological termsM. If metatheory is taught Lin such an ideological wayM (p.../), then, Linstead of opening up for thin!ing, it closes down the path to debateM%

If diverging values are all there is, then the debate can easily run into a show of verbal fists. 7orse, in some settings, the intellectual e\$change might never start since everybody feels entitled to stic! to what they thin! anyway (and professors are always right).

In other words, despite whatever positive roles that metatheory can potentially play in education, they may also end up being, Lparado\$ically, undermined by a lac! of reflectionM (1ates and 4en!ins 5//>a, p.=.).

+ne last point connected to teaching is that the advantages of learning metatheory, as seen above, are usually depicted primarily in terms of transferable skills such as Kencouraging critical thin!ingP. 6olin -ay (5//>, p...A), one of the te\$tboo! authors discussed by 1ates and 4en!ins, portrays their 9ustification for teaching ontology and epistemology in political analysis as Lcouched almost entirely in terms of the indirect rather than the direct conse"uences and benefits of being able to ad9udicate argumentsM. +pposing the primacy of such indirect benefits in 9ustifying the educational use of metatheory, -ay (5//>, p...=) states that Lit is wrong to suggest that these are responsible for its place in the teaching of political analysis todayM.

(oreover, while none of the pedagogic KgoodsP to which 1ates and 4en!ins appeal can be imparted directly, none can *only* be attained through a consideration of ontology and epistemology. &he relationship between the goods to which they appeal and the means by which such goods are to be attained is, then, almost entirely contingent. Indeed, there are clearly other 0 and, arguably, more effective 0 means of promoting each of the benefits to which they refer (p..==, original emphasis).

For this particular reason, if transferable skills are the main justification for teaching metatheory, then the whole enterprise may be jeopardised. Those less wedded to the advantages of ontological and epistemological reflection than Latour and Woolgar might well use precisely the same move to argue for the demotion of such meta-theoretical debates from the curriculum. To sum up, metatheory has sporadically been defended as a good didactic and emancipatory educational device. However, its didactic value is highly contingent upon using adequate teaching methods and positive results are less than likely to obtain. From an emancipatory perspective, in turn, metatheory is said to lead primarily to certain benefits connected to transferable skills. Despite this, results are again less than guaranteed and, besides, several other kinds of content could be taught instead in order to develop the same mentoring-related skills. If there is any advantage in teaching metatheory to politics and IR students, it must be defended on other grounds.

The politics of metatheory

Metatheory is said to play the negative roles of bringing further complexity to abstract research and its communication in the disciplinary context of IR. As seen before, this may be due to both intrinsic and extrinsic features of metatheoretical knowledge. Apparently, the same scheme applies to the assessment of political implications of metatheory. Some scholars have protested against the increasing 'metatheorisation' of IR by associating it, first of all, with the negative inherent political effects of the conflation of distinct discursive orders and practices. Secondly, it has also been pointed out that metatheory is a hindrance to the pursuit of the social calling of intellectual life. Finally, some have claimed that the metatheoretical dimension of theoretical thought operates as a mask for all sorts of ideological struggle and other political issues related to social scientific scholarship. These objections to metatheory come from varied sources, but it is clear that they draw on previous commitments to specific positions about, for example, the role of IR or social research and the role of intellectuals in society. It is, therefore, important to turn to each of these points in more detail by also taking into account their backgrounds.

First of all, then, there is the notion that metatheory plays a negative political role due to some of its inherent discursive-order features. This idea is yet another application of Tunnell's (1999) claim that metatheory, for most part of the time,

conflates the first (social practice), second (theory about social practice) and third orders (theory about theory of social practice). -is intention is to provide La criticism of (...) the assumption that academic discourse is necessarily in some way a form of political actionM (p.=), an assumption which, according to him, is increasingly recurrent in politics and social science specifically due to the rise in popularity of poststructuralist and postmodernist approaches. &his is highly problematic because first'order practices (for instance, world politics) are simply Lnot normally or logically open to metatheoretical arbitrationM (p.5/). Why, then, is there an insistence in treating first' order practices as sub9ect to higher discursive orders8 &he problem of KmetatheorisingP the study of first'order practices derives, according to Zunnell (.99@, p.5=), from its Ldual rhetoric of in"uiryM, that is to say, a strategy Lwhich is devoted not only to the internal 9ustification of certain claims and approaches but to 9ustifying itself to an e\$ternal audience that is often its ob9ect of in"uiryM. In IR such strategy may be illustrated by 6hernoffPs (5//>) study of policy'related implications of metatheory. -e attempts to e\$plain why practitioners should employ metatheoretical principles in order to test competing IR claims before adopting specific policy recommendations emerging from them. &he purpose of this discursive approach is, according to (onteiro and Ruby (5//9, p.GF), to show that L!nowledge produced by a science of IR that stands on un"uestionable foundations can claim the status of a special discourse about international relations, one that cannot be challengedM. ,or Lwithout establishing their enterprise as KscientificP, how can IR scholars e\$pect policyma!ers to tale the intellectual fruits of the field seriously8M (p.=) &herefore, in the same way that metatheory must not be conflated with theory, it also should not be employed as foundational for political practice. &he fact that it is reflects an an\$xiety related to the status of the discipline rather than any logical necessity.

&he conflation of discursive orders and practices can, alternatively, be seen as a move aimed at alleviating the political pressure from scholarsP shoulders. 4ames <urthPs (.99@, pp.F5'F, my emphasis) opinion voices the general sense of the policy community that too much abstract research reflects an isolationist tendency in IR scholarship. &he negative effect is that, instead of being concerned with pressing political issues, scholars turn their focus to their own academic peers and, thus, abstraction turns into an end in itself%

(most academics (...) are only academics (...) X and Y one-dimensional people. & they therefore have not had the breadth of experience to give them a sense of the reality of international relations; nor have they had the depth of experience to give them a sense of its tragedy X and urgency Y (...). For such people, there are only a few kinds of achievements possible in regard to international relations. One of them is to sit down and spin out theories that fellow academics will praise as being rigorous and original. (Merely being *realistic and responsive*, in contrast, will not evoke their praise). In truth, most academics are only concerned about the good opinion of about a dozen other academic specialists in their particular sub-field.

Obviously all the accusations above potentially apply to every kind of theoretical scholarship. As a foreign policy theorist once put it%

I am afraid that even some of my academic colleagues exaggerate the promise of theory to the point of making theorising not an aid but a substitute for really hard thinking about the issues on hand; it is much cosier to delve into the intricacies of an abstract theory than to analyse, leave alone resolve a concrete issue (Ran!el, 1990, p. 100).

However, Kurth (1990, p. 100) has IR *metatheory* specifically in mind. & the rise of metatheoretical debates, for him, means that in the future, the study of international relations will be even more uninteresting and irrelevant than it is now. Commenting on the loss of political content with the increasing philosophicalisation of IR, Monteiro and Ruby (1999, p. 100 note 10) summarise the issue%

Perhaps this helps explain why foundational questions are particularly attractive to IR scholars. & the real-world stakes of academic IR debates can be very high, up to and including questions of life and death, or war and peace. So IR scholars may succumb to the temptation to settle political debates on big issues of international relations by invoking the status of an unquestionable discourse that of a science with unshakable foundations.

In both cases too much theory and too much metatheory & the avoidance of politically pertinent problems is ultimately a negative outcome of discursive order conflation.

& the distance, or gap (Bye, 1990; Walt, 1990, 1991) between (meta)theory and political practice is particularly acute in relation to the perceived mission of social and political science in society. Since their inception, a sense of calling or vocation has certainly been present in academic economics and sociology (3. & 4. Jackson, 1990). Political analysis, for that matter, developed itself against a civic educational background and, at least at an earlier stage, internationalist academic discourse followed similar tendencies by emphatically stressing its own role in democratic society (Schmidt, 1990, pp. 100-101). (More specifically, we are often reminded that the academic institutionalisation of the study of world politics initially aimed at nothing less than the

eradication of warfare (Ritt, 1954; Nuttall, 1990, pp. 5-15). The rationale behind having a separate discipline of IR, as one of its pioneers once put it with some exaggeration, was that the better the world is understood by the better people in it, the better for the world will it be (Manning, 1949, p. 10). Since the beginning, therefore, there has been a working assumption that the discipline should pursue the improvement of the general conditions of coexistence in the international scenario, given the pressing issues of political practice (see also Ritt, 1954). Nicholson (1999, p. 10), whose position is overall in favour of metatheory, voices the same concern for the contemporary setting. He also believes that the political relevance of IR goes beyond the political relevance of performing civic duties.

The purpose of doing International Relations, like all social science, is to influence people, sometime, somewhere in a context which will make a difference to their actions. Thus, at some stage, possibly distant, a course of action will be taken, or abandoned, as a result of our efforts. The world will then look slightly or even significantly different because of our activities. We hope it will look better, though what Ketterer consists of is itself a result of our moral positions. This also means that we think we have something to say which goes beyond that which a concerned citizen could say.

Arguments about political relevance of scholarship are, in short, usually made from general assumptions about the role of IR in society.

This provides the background for the second point: metatheory is also portrayed as playing a negative role with respect to the political implications of social scientific study. Perhaps the position most often associated to this theme is that of critical theorists, who have connected the role of IR scholarship to emancipation, at times, via sociological imagination. The notion was classically defined by Wright Mills (1949, pp. 1-2) as a quality of mind essential to grasp the interplay of man and society, of biography and history, of self and world and that enables one to use information and to develop reason in order to achieve lucid summations of what is going on in the world. Sociological imagination has been advocated as defining IR's vocation. It is, as Austin Rosenberg (1996, p. 10) argues, a necessary *public* virtue. Unless we know what the structural consequences of our choices are, we scholars cannot exercise the political freedoms we have in a responsible way. He claims that sociological imagination as applied to IR must also reflect grand theory and excessive abstraction, even in empirical research, and that this is paramount to emancipation.

It is (O) the particular responsibility of political intellectuals to promote (O) collective self-knowledge by illuminating the structures of social power, identifying the kinds of human freedom which can be realised within them, and tracing where the lines of responsibility (O) may run (O). X&his responsibilityY is addressed to a real social need, one which the alternative vocations of grand theory and abstracted empiricism have disavowed themselves from answering (O). &his need is surely no less pressing in International Relations than elsewhere.

The problem of IR is that it has crystallised into a bureaucratic social science by distancing itself from classical (i.e., Wright (ill)Ps) social analysis. The grounding in substantive problems, the use of an historical and comparative depth of field, the perception of social totalities, and the commitment to ideals of freedom and reason as a way to pursue the duty of sociological imagination (1996, p.96). In short, the shift from substantive issues and a concretely historical view of things, as well as the escape to technicalities of metatheory, have been portrayed as a deviation from the vocation of IR as a discipline which is supposed to be socially and critically relevant.

Concern with the social role of IR scholarship is not the monopoly of critical theorists. Peeling from the mainstream, Han &vera (1997, p.4, my emphasis) discusses the duty researchers have towards the general community as a matter of professional ethics.

Social science operates largely beyond accountability to others. Institutions and professions that face weak accountability need inner ethical rudders that define their obligations in order to say (*sic*) on course. Otherwise they risk *straying into parasitic disutility*. Social science is no exception.

Lacking established social relevance as part of the professional ethics which politics and IR scholars are supposed to follow, Han &veraPs (1997, p.5) distaste for metatheory comes as no surprise. Later on, in a list of principles that purports to define what an interesting research topic looks like, he recommends caution in the selection of highly abstract questions, as the study of the fundamentally unknowable is futile and should be avoided (p.6). By implication, this means that IR researchers have a duty to address only those problems which are regarded as pertinent according to his criteria.

We have an implicit contract with society in exchange for academic freedoms and privileges we agree to spend at least some energy answering societyPs more urgent questions. &his does not require headline chasing. We can meet our obligation with policy research or with more abstract work that could have policy implications far down the road. But social science violates its contract if it drifts into complete irrelevance, as much of it has (p...@).^{F9}

F9 Even in many of the critiques of abstract research, there is no further explanation of how we could distinguish more abstract work that could have policy implications far down the road from work

Wallace (1998, p.14) concurs by stating that Lwe owe a duty of constructive and open criticism% to speak truth to power, and not (O) to speak truth in secret only to each other. -e provides a clearly negative diagnosis of British IR and blames excessive abstraction as the cause of its deviation from such political duty. The discipline Lhas become too detached from the world of practice, too fond of theory (and meta'theory) as opposed to empirical research, too self-indulgent, and in some cases too self-righteous (Wallace 1998, p.14). -olsti (1998, p.14) affirms that, in IR, Lthe tensions between theorists and practitioners are well known. The gap which time and again emerges between practice and scholarship are at least partially associated to metatheory.

suffice it to say that communications between them ends when the one is unable to make any sweeping generalizations from vast experience in the field, and the other engages primarily in epistemological and lofty metatheoretical argument. The one drowns in a sea of hyperfactualism, the other in an ocean of metaphysics.

It is clear, then, that metatheory is perceived (by scholars coming from different places across the disciplinary spectrum) as playing the negative role of insulating IR from practical engagement.

The fact that metatheory implies distance from concrete politics and from the (perceived) role of scholarship in society may also go together with a third political problem% metatheoretical discourse has more than once been criticised as a technical mas! for highly politicised issues.

The isyphean effort to place IR on unshakable X3hilosophy of scienceY foundations is at least in part an attempt to boil down the complex relation between IR and international relations to the supposedly more straightforward relation between social science and philosophy. In other words, it is an attempt to turn a political problem into a philosophical:foundational one (O). This reduction, however, is an over'simplification. There is no way to strip the relation between IR and international relations from its political dimensions. There is no coolie cutter way to deal conclusively with this complex issue in an apolitical way 0 a "uestionable goal in itself ((onteiro and Ruby 1998, p.14).

One of the aspects of scholarship hidden by metatheory, some claim, is the fact that academic discourse is inevitably driven by ideological assumptions. By making the whole discussion look purely technical, scholars attempt to camouflage the political roots of their argumentation. An example of this argument may be found in Andrew

that Ldrifts into complete irrelevance. Is this not, as Ealatos (1998) has extensively argued, a judgement more properly issued by hindsight?

Lawrence (1999, p.99) critique of the Imperial method employed by democratic peace researchers. The author is not explicitly against metatheory and actually employs it in order to develop his claim. The idea is that democratic peace research has been treating political problems like the definition of 'democracy' at the (wrong) level of metatheory. By connecting positivist notions of scientific research to IR theory via a restrictive set of methods, theorists arrive at a superficial notion of 'democracy'. However, in order to establish the thesis that democracies do not fight each other, the very definition of democracy is at stake. Dealing with such an important issue at a non-political level leads democratic peace theorists to beg the question.

Referring to metatheoretical debates in IR over the notion of science, (1999a, p.100) argues that Ignoring the politics of 'science', philosophy of science runs the risk of misunderstanding why 'science' debates matter in IR and what is at stake in them. The line between politics and metatheory exists, according to her, not because a specific metatheoretical position is strongly determined by a particular political ideology, but rather because political worldviews are 'constraining and enabling' (p.100). For example, someone who favours top-down social engineering is highly unlikely to support such political preference from the point of view of a poststructuralist, anti-closure, approach. As a result, metatheoretical issues must not be treated as merely philosophical and abstract, lest we run the risk of assessing the situation by ignoring such a relevant aspect as politics. She complains about the current tendency to engage in a curious pretence of apoliticality in 'science' debates in IR (p.100) and concludes that 'we should at least remain open to the proposition that 'philosophy of science' positions can have certain political predilections built into them' (p.100). In other words, it becomes impossible to maintain that metatheorising is devoid of any political character, despite the current tendency to treat metatheory from a neutral perspective. This last point closes the circle of negative perspectives on the role of metatheory in IR: it leads to an increasing insulation of the field from political practice, it prevents a higher level of critical engagement in society and it disguises certain politicised features under the mask of technicality.

Discussion

Several negative views of metatheory in IR have been presented. In this examination of the relevant literature on the topic, I have identified four key areas in which metatheoretical research is said to hinder the development of the discipline. Due to its

intrinsic features, to the complex character of the issues that pertain to the subject, to educational aspects and, finally, to political impact, metatheory has been perceived as not playing a constructive role in IR. By voicing these concerns, I have gone beyond an application of the previous chapters (i.e., how metatheory operates in general) and towards an examination of some of the negative externalities which have emerged with an increase in metatheoretical activity in this specific field. I turn, now, to a brief critical examination of each of the points presented in this chapter. Leaving a more detailed critical procedure for later, I focus here on how these elements hang together, how some actually cancel each other out and how, individually, some of them are internally problematic. We should take IRPs' critique of metatheoretical research seriously, but the strong programme of eliminating it from the discipline is untenable.

So, first of all, what are the main relations between each of the complaints raised against metatheory? Let us recall the analytical objection to metatheory based on its intrinsic features of a higher-order discourse as a key converging point. This critique of metatheoretical research implies that it must necessarily be differentiated, on the one hand, from theory and, on the other, from social practice. IR, as a study of social practice, must not be conflated with the study of such study (i.e., a meta-study). If, to the contrary, such conflation does occur (as it is apparently the case in the discipline), then we will be operating under the illusion that a meta-study is capable of dealing with a lower-order practice. For example, searching for higher-order criteria in order to assess lower-order features is said to be a misguided strategy that leads to negative consequences for a field of inquiry, such as forgetting about a theory's own merits and judging it by higher-order criteria (e.g. philosophy of science) or, for that matter, presenting metatheory as extremely relevant to political practice in order to justify our production of knowledge in the discipline. In these two cases, metatheory plays negative roles such as increasing the complexity and fragmentation of IR along higher-order discourses (and not the directly relevant theoretical order), and attempting to justify IR via its higher-order status instead of making the case for its direct political pertinence in practice. Such potential consequences of metatheoretical activity cannot in any way be dissociated from its Intellectual or intrinsic discursive-order features. Therefore, there is a relation between the intrinsic features of metatheory and its complexity and political roles.

Particularly relevant too is the connection between the increased complexities of IR due to the metatheoretical tendency to conflate discursive orders and some of the

issues of communication. Because it fragments the discipline along higher-order divides, metatheory is at least partly responsible for the current lack of understanding across approaches. Insisting on judging theories by their metatheoretical leanings, IR scholars reproduce and amplify divides to such an extent that it becomes very hard to bridge them back to synthesis. This is perceived as leading to an aggravated isolation between research communities in the field. Another aspect of such isolation comes from the fact that metatheory also imports knowledge from other disciplines, often with much delay and decontextualisation, which renders this discursive layer even more problematic in terms of communication. This being the case, it comes as no surprise that, besides the cross-approach divides, another distinction emerges between those interested in more abstract research and those who are not. It is also not surprising that metatheory may be regarded as an excruciating intellectual activity% we often distance ourselves from that which we do not understand. Moreover, this kind of distinction that emerges from communication problems is not a trouble faced solely by scholars. In fact, if the research community itself cannot get past this issue, then why should we expect students and other outsiders such as practitioners and the media to understand what is going on? It is also not surprising that an aversion towards metatheory has developed amongst those outside the field who are, nevertheless, interested in understanding international political practice. In addition to this, we even run the risk that, in our attempts to make metatheory look simple due to such communication problems, we may actually end up making it look simplistic. And with false simplicity comes a false assurance that we have arrived at a final answer to these questions. This kind of closure is especially damaging in the case of teaching, where closure to some extent may create a path-dependent decrease in theoretical plurality in the discipline.

2. Deepening in teaching, disciplinary enclaves, bureaucratic insulation 0 all such elements reflect a theme that has been so far ignored by the literature if compared to the other points. Metatheory plays a negative political role not only in relation to practitioners or, perhaps, to others who want relevant scholarship in order to change society. What this is the case is sufficiently clear from what we have seen. However, the negative political externalities of metatheoretical research must also be seen in terms of the internal disciplinary politics. It is with respect to the politics of scholarship that the further political problems originate and spread. Hence, for example, Cynthia Weber's (1996, p. 11) portrayal of Chomsky's (1989) attempt to discipline IR feminism in political terms, as yet another example of male paranoia. Perhaps it could be argued

that even the (metatheoretical) critique of metatheory based on political considerations serves the purpose of masking the politics of metatheory whenever it ignores the disciplinary political aspects. There is indeed a relation between metatheory and the fact that some of the politics behind the debates is somehow concealed. Nevertheless, the notion that this goes beyond the usual connection between our inherent political involvement as social actors and the accident that some of us happen to be scholars too must be taken to a further step. It must be developed into an analysis of the politics of metatheory concerned with its intra-disciplinary political effects. This leads me to consider the next issue at hand.

Besides looking at how the points raised about the negative roles played by metatheory in IR converge, it is also necessary to take a critical look at how they diverge. In fact, some of them cancel each other out. Still on the subject of politics, for example, it is clear that two mutually exclusive theses have been advanced. One is that metatheory isolates IR from politics, thus leading to a decrease in the field's relevance. The other is that metatheory actually brings politics to IR in particularly damaging ways such as, for example, serving as a kind of Trojan horse for the clash of political ideologies. Both claims cannot be sustained at the same time and in the same relation. If one is to take issue against metatheory, then the option must be made between one point and the other. The same applies to the argument related to the complexities brought to IR by metatheory, some of which rely entirely upon the assumption that metatheory furthers the depth of the disciplinary divides. If that is the case, and if insulation is really a negative configuration, then why should we attempt to amplify fragmentation even more by drawing a sharper line between metatheoretical research and the rest? It seems that one side of the argument calls for more dialogue and plurality in the disciplinary discourse, whereas the other attempts to drag metatheoretical research out of the field by crystallising the distinction between discursive orders into a rigid institutional framework, thus leading to even more fragmentation. A similar contrast operates in the case of communication. If metatheory is to be further isolated from the rest of the discipline, how can we expect it to become more intelligible and perhaps more interesting? (Maybe it is the opposite, i.e., bringing metatheoretical research to *further* integration with IR, that will help deal with the issue of communication and, by extension, with the educational aspect.

Going beyond comparison and contrast between the negative points concerning the role of metatheory in IR, it also must be said that some of them are internally

problematic. The discursive order argument, for example, which is crucial for establishing a strong opposition to metatheory in any given field, depends on a falsely characterised situation. It may be the case that theories should be judged by not only metatheoretical criteria, but also other notions that do not pertain to this discursive order. However, we would go too far if we denied metatheory any relevant role based on the wrong premise that a discursive order does not have any impact or authority over another. In the example presented before, the claim would be that texts about football do not have any impact on football playing, which is obviously wrong. Of course, not all texts about football will have such impact, but some will. It is relatively easy to conceive of a situation in which, for instance, a manager will read journalistic commentary about how her team is doing and make decisions based on that second order discourse. By the same token, IR theory could have a certain impact and even be regarded as possessing a certain authority over policymaking (even if empirically this does not seem to be the case), in the same way that metatheoretical works may be read by IR scholars and somehow orient what they do. It is precisely in those terms that Winiarski (1999) portrays the rise of IR constructivism, a theoretical approach born out of metatheoretical considerations. In addition to this, metatheory may also play a part in theorising in that, for interpretivist scholars, theories are embedded in the social world that we study. Anarchy is what states make of it (Wendt, 1995, p. 99A) because states 'believe in' certain theories of anarchy. Lower politics became the prevailing mode of action in the modern international system because, first, it became the prevailing mode of thought (Wendt, 1995, p. 99B). Metatheory is, therefore, also relevant at the *theoretical* level (see also Winiarski 1999A). Actually, the fact that *it is* has led IR to proliferate its metatheoretical literature. The existence of problems with metatheory actually testifies against the strong argument in favour of eliminating it from the discipline. Not only this, but the productivity of metatheory in terms of leading to new approaches at the lower, theoretical, level illustrates that metatheory in fact is not merely parasitic upon theory, but can be 'productive of' theory (Ritzer, 1995, p. ...).

We are, thus, left with a weak version of the claim instead of discussing whether or not there should be any metatheory in IR, we are left with the question of whether there is too much metatheory in IR. Accordingly, it may be granted that metatheory alone cannot judge political or theoretical arguments, that metatheory may cause misunderstanding, that metatheory may be hard to teach and that it may be hard to enjoy. All these arguments are perfectly valid in their context but they also possess quite

a few weakening properties if we attempt to extend them. The claim that metatheory is uninteresting is purely based on subjective opinion. It could be the opinion of many, but it is most definitely not the opinion of all. (Moreover, if the quality of research and communication of metatheoretical knowledge has been low, it does not mean this will always be the case and, besides, it is just not clear how less incentive to metatheory would contribute to improve the situation. This does not mean that everybody should be doing it. In fact, there is a certain agreement among IR metatheorists that the relevance and use of metatheory is always conditional upon being a tool for something else and not an end in itself (e.g. Allier, 1997; Beufeld, 199A). These views are echoed in -olsti's personal assessment of the metatheoretical debates%

I am somewhat concerned that too many people may be spending time discussing great issues of epistemology and metaphysics. I am interested in international relations more than philosophy. We are here because we are interested in a field of human activity. We must be sensitive to the ways we approach its study and raise questions about the whys and hows of knowledge. But beyond a certain point and I cannot define exactly where it is located concern with epistemology may lead us to lose sight of the subject matter (interviewed by Jones 5/5, p.5F).

Else -olsti, Brown (5/5, pp.G.A'G.=) recognises the partial merits of metatheory as well as the need of cultivating a balanced view of the issue%

That we have here are sophisticated and intelligent contributions to debates in the epistemology and ontology of the social sciences, but it is much less clear (at least to me) that these are debates that ordinary practising social sciences need to get too worried up about.

(Metatheory may play several negative roles, but neither of them provides a compelling case for eliminating metatheory from IR. On the other hand, metatheoretical research should not be the centre of every scholar's attention in the discipline.

Final Remarks

In this chapter, I have considered intrinsic and contingent negative roles ascribed to metatheory by the IR community. Having expounded them in detail and very briefly analysed their merits, I come to the conclusion that the radical case against metatheory fails on account of internal problems. (Moreover, weaker objections to metatheoretical research are usually associated to historically accidental features of the development of this discursive layer in IR. This being the case, it does not follow that we should eradicate metatheory from the discipline. However, it also does not necessarily follow

that we should do more metatheory than we do now. What has been established so far is simply that% (.) the status of metatheory in IR is highly contested; (5) several negative outcomes have emerged from metatheoretical research in the field; (F) none of these outcomes poses a sufficient reason either for interrupting our engagement with metatheory or for increasing it; (G) a number of the arguments raised against metatheory cancel each other out; (A) some of the claims raised against metatheory face considerable problems in the way they are structured; (=) in some of the negative views of metatheoretical research, its eternalities are weighed against the positive potential of metatheory. This last point leads me to the next chapter% if there are some positive expectations about metatheoretical research, we should look at the extent to which they are actualised in the academic discipline of IR.

Chapter .

Positive Views of Metatheory in International Relations

Introduction

Having discussed the main complaints against metatheoretical research and debates in IR, I now turn to perceived positive roles of metatheory in the field. The method is still the same as before% by critically reviewing the relevant literature, the main goal of this chapter is to present several views of what metatheory does in IR. This time, however, the focus is on the perceived actual and potential positive contributions of metatheoretical scholarship. The main source to be used here is the body of texts mainly concerned with IR theory and other self-referent disciplinary issues. Authors dealing with these research topics often make side remarks on the relevance of this kind of study, and asking how they justify the potential contributions of their own metatheoretical research yields insight into their perception of the positive roles of metatheory. Moreover they may also anticipate anti-metatheoretical criticism and attempt to explain the worth of their metatheoretical enterprise (or even metatheory in general) in reply. The second main source of positive comments on metatheory in IR is a collection of scattered comments in predominantly non-metatheoretical texts. Review passages in textbooks, theory textbooks and pieces on disciplinary history are the most likely elements to fall under this category. Finally, a third important source derives from texts which do not reflect much self-awareness of their metatheoretical character, but which may, nevertheless, be classified as predominantly metatheoretical according to the definitions provided and expanded in this thesis.

While I focus again mainly on the IR literature, it is nevertheless crucial to understand that many of the claims derive from neighbouring disciplines. For this reason, going beyond the IR literature wherever there is a clear connection between intra-IR claims and extra-IR discourse is a welcome additional procedure. The main headings for the positive views of metatheory are% (.) the potential improvement of IR theory and; (5) a better understanding of the discipline, including (but not restricted to) a refined analysis of its historical formation; (F) clarification of theoretical material and

its assumptions; and (G) study of theories as part of the social world. In my discussion it becomes clear that most of the points in favour of metatheory contradict the main negative views in IR. As a result, besides their mutual cancellation in some cases (as seen in the previous chapter), the complaints against metatheory also face the challenge of the positive views based on similar (sometimes exactly the same) issues. We are left, thus, with the following conclusion: if a certain range of points indicate both constructive and negative roles of metatheory in IR, then each of them will have to be assessed on an individual basis. In other words, there ceases to be a *prima facie* case against metatheory. However, an even stronger response to the complaints against metatheory in IR can be made. At the end of this chapter I provide an additional analytic argument for the validity of metatheoretical enterprise that problematises any attempts to completely eliminate metatheory from our academic field.

A tool for improving IR theory

There are several ways in which metatheory is perceived as playing a positive role in its contribution towards the improvement of IR theory. Contrary to the claim that metatheory is merely parasitic and does not lead to new theory, (Ilina & Ullrich (1995, p.5.F) argues that metatheoretical framings of explanatory frameworks have direct effects on the kinds of explanations we advance for concrete world political processes: indeed, theoretical and conceptual lenses constrain and enable the kinds of explanations we can construct. One of the key points made in favour of metatheorising is that it may lead to better theories of world politics and, in fact, in concrete cases it has actually led to new theoretical material which has gained wide acceptance, such as neorealism and constructivism. Another argument for metatheory is that part of it consists in evaluating and validating research and this, in turn, refines IR scholarship. (Metatheory, therefore, would be a necessary means to that end. A third way in which IR scholars portray metatheory as theoretically productive is the claim that a better organisation of scattered theoretical material leads to new theories. Most of the key IR theorists (e.g. Morgenthau, Hall, Waltz, Wendt, Kohane) have, at one point or another, dealt with metatheory in a theoretically productive way. Viewed in this way, says Joseph Aupille (1994, p.5.), metatheory sets forth the basic architecture and requirements of scientific research, both guiding it and providing standards by which it can be assessed by members of the scholarly community. The general agreement is that, ultimately, by doing this kind of metatheory, we aim at developing a more reliable

and systematic understanding of international relations (Eieber, 1997, p.viii).

A well-known example of IR theorist dealing productively with both theory and metatheory is Kenneth Waltz, whose key works, *Man, the State and War* (1959) and *Theory of International Politics* (1979), are among the most cited books in the discipline. Elsewhere, I have provided a detailed reading of Waltz's work as a coherent polyphonic discourse of three analytic voices: a metatheoretical voice, a disciplinary historical voice and a theoretical voice (E. 2. 1999, 5//b). Waltz's *modus operandi* throughout his career is a good illustration of the three points mentioned above: a metatheory as collecting and organising material for new theories; as research evaluation/validation and as enabling new theories to emerge. The initial move is primarily contained in the 1959 volume. The book gathers theoretical material on the causes of war from sources as varied as classical political theory, memoirs of diplomats and state leaders and the discipline of psychology in order to organise them into coherent sets of images or categories of explanation. The final section presents a critique of each image and suggests the next steps of research, hinting at the desirability of placing all images together in an overarching framework. We learn from Waltz (1979, p.F9.) that in the 1950s Waltz decided to spend a few years reading on theories and the philosophy of science. In 1979 he was able to introduce neorealism in IR as emerging from his metatheoretical views.

I write this book with three aims in mind: first, to examine theories of international politics (O); second, to construct a theory of international politics that remedies the defects of present theories; and third, to examine some applications of the theory constructed. The required preliminary to the accomplishment of these tasks is to say what theories are and to state the requirements for testing them (Waltz, 1979, p., emphasis added).

In the *Theory* there are two long metatheoretical sections: one on what theories are and how we should test them according to philosophical criteria; and the other is a concrete assessment of past IR theories according to Waltz's metatheoretically adjusted set of criteria for the discipline.^{G/} In his own words, metatheory was 'preliminary' to the construction of what he believed to be a better theory of international politics.

Not only has metatheory been crucial in the elaboration of theoretical material (and explicitly seen as such) in Waltz's case, it also has become paramount to further

^{G/} That these are two distinct ways in which metatheory contributes to Waltz's formulation is clear enough from the fact that the sets of criteria employed in each do not necessarily overlap. Chapter 1 of this thesis includes a discussion of Waltz's 'instrumentalist' view of theories and Chapter 2 mentions his preference for 'systemic' over 'reductionist' theorising.

(meta)theorising in response to neorealism.

scholars who had so laboriously opened IR to the new frontiers during the 1990s. The interparadigmatic debate may have to watch how the success of Walt's *Theory of International Politics* locked up the discipline again (O). &he critique therefore aimed at the meta-theoretical foundations of Walt's new disciplinary wall. Indeed, one can say that Walt's critiques (*sic*) used the debate around his book to force an inroad of meta-theoretical discussions into IR (Zürn, 2003, pp. 10-11).⁶

A particularly relevant post-Waltian IR approach that clearly emerged from metatheoretical considerations is mainstream constructivism, as represented in Alexander Wendt's *Social Theory of International Politics* (1999). (More than one commentator of constructivist research shares this reading. (John Ruggie's (2000, p. 119) narrative reconstructs the prelude of Wendt's intellectual journey this way%

Firstly, he introduced the agency/structure *pro lematique* within the theory of international relations. &hat was a clarifying as much as a complicating move in the meta-theory of International Relations (O), in any case important because of his persuasive introduction of scientific realism and its possible uses in social and international theory (O). Secondly, in perhaps his most commonly "quoted article, a new agenda of social constructivism was accepted and developed in a more scientific way (O) than the original XconstructivismY, which introduced (O) the term itself to IR.

Stefano Zürn's (2003, p. 10) account also locates the starting point for IR constructivism in a 'meta-theoretical shift', a unique moment in disciplinary history. Whereas previous schools of similar content fought their battles on the level of policy analysis, this time the struggle was drawn into the meta-theoretical field. Constructivism combines many old hats with a willingness to challenge the scientific project of (mainstream International Relations) (2003, pp. 10-11). &he cases of neorealism and constructivism are, then, two examples of how metatheoretical considerations have actually constrained and enabled substantive theoretical formulations.

On the historical side, besides Walt's first book, there are other examples of the harvesting and organising role of metatheory directed particularly to the elaboration of contemporary theories from material which might be available, but in an unsystematic, incomplete and scattered way. A great deal of (artin Wright's published work consists exactly in this kind of metatheoretical exercise. James Cerco (1999, pp. 10-11), Roger Handberg (1999) and Ian Hall (2000) correctly see Wright's emphasis on theoretical and

G. Hall (2000, p. 119) comments that later on, when Walt was attacked for postulating at the foundation of his theory some axioms which were unfaithful to reality, his reading of the scientific method was "quite handy".

political discourse as a way of studying *international society* itself (see, *ibid.*, 5/5). In this sense, metatheory operates as a tool of theory.^{G5} Wright (1999) was also concerned with the status of IR as a discipline with no coherent theoretical body despite the wide availability of practical discourse about world politics. The international theorist's role, therefore, included the gathering, exposition, critique and refinement of this material. The notion of harvesting and organising apparently disconnected propositions as a step to new theory has been a recurrent justification for works operating in the intersection between political theory and IR theory in particular. Thus, for example, we read in Nuttall (1999, pp. 15) detailed testimony on the history of international theory that he wants to counter the assumption that there is no IR theory outside IR scholarship by stressing the relevance of pre-disciplinary international political thought. The point here is that sometimes we do not need to reinvent the wheel and that, after we organise them, these already available ideas will actually provide us with the necessary building blocks of new theory.

Michael Conelan (1995, p.), in his preface to an interesting imaginary dialogue between representatives of different reconstructed traditions of international theory, states that he aims to 'take the basic thoughts of the great men and try to widen them out to international politics'. However, he makes it clear that our aim is not to describe past thought, but to use it to build theory (1995, p. 5). A similar argument appears in Zallie (1999, p.) monograph on *Philosophers of Peace and War*. The metatheoretical character of the book transpires in his statement that it is meant 'for comparison, analysis and assessment' of theoretical material which is scattered because the authors form 'not a school, nor even a clear succession or progression of thought'. Alluding to Wright's approach, the expressed reason is that there is a need 'to collect our thoughts, to begin to unify our still separate lines of thinking on the issue'. Zallie (1999, p. 5) then adds an interesting twist to the harvesting and organising function. 'To those willing to understand the complexities of war and politics in the nuclear age, there should be an interest in theory formulated at a time when it seemed very much simpler than we know it to be'. It might be an exaggeration, but the claim that 'the more we appreciate the achievements of past thinkers' (O), the better equipped we should be to deal with the new complexities is clearly a defence of metatheoretical history of thought leading to new thinking on contemporary issues. Showing the possibility of a similar use for metatheory when we understand ideas about world politics coming *from* IR itself, Brian

G5 I further develop the point below in this chapter.

Schmidt (1995, p. 6, emphasis added) summarises the point by listing some of the key aims of disciplinary historiography:

First, numerous theoretical insights' of largely forgotten scholars' have been simply erased from memory! Even once recalled' these insights can have critical purchase in the present. Second, the field has created its own powerful myths regarding the evolution of the field that have obscured history (O). Third, an adequate understanding of the history of the field is essential for explaining the character of many of our present assumptions and ideas about the study of international politics (O). There is, in fact, much evidence to support the proposition that much of what is taken to be new is actually deeply embedded in the discursive past of the field. Finally, a perspicacious history of the field offers a fruitful basis for critical reflection on the present (O) and for opening up some much needed space in which to think about international politics in the new millennium.

Other disciplinary historians would confirm the point (Higneswaran D Nuir, 1995, p. 17).

A tool for understanding the IR discipline

Metatheory has also been portrayed as essential to a critical understanding of IR as a discipline. There are three fundamental ways in which this is perceived to be the case, and much of current metatheoretical research actually falls into these categories: the provision of a self-image (and critique thereof) in IR, the specification of our object(s) of study in the field at the most basic level, and the analysis of limitations and challenges in the discipline.

The abundant elaboration and critique of disciplinary self-images in IR hints at dissatisfaction with some of these self-images and, in some cases, with metatheoretical self-imagetic discourse as such (e.g. George, 1999). However, metatheory is still defended as the forum in which these self-images and the roles they play are critically evaluated (Smith, 1999A). Well-known examples of highly contested self-images are: the Science:Art dichotomy (e.g. Beal & Hamlett, 1999); the KIR as an American social science portrait (e.g. Hoffmann, 1999); the K systemic:Reductionist distinction (e.g. Walt, 1999); the Kreat Cebates narrative in IR's disciplinary historiography (e.g. Janis, 1999A); the K traditions of international thought narrative ((Wright, 1999); the organisation of the field around ontological (e.g. Uri, 1999; Wright, 1999) and epistemological (e.g. Hollis & Smith, 1999; Abri & Khan, 1999) categories, among others. In several of the cases above, a new metatheoretical self-image has been produced exactly in order to correct problems in a previously attained self-image, such

as the lack of debate across divides due to a misapprehension of where the KrealP divides are (Jackson, 2005, p.17). Disciplinary history is prominent among many relevant kinds of metatheoretical research in this process. According to one of its key proponents, its aim is to problematise the prevalent interpretations of how the field has developed and to indicate that the history of the field is both more complicated and less well known than typically portrayed (Schmidt 2005, p.6). As a tool for understanding IR, then, metatheory goes beyond the formulation of a systematic reading of the field; it is also the space for "questioning such metatheoretical readings.

An illustration connected to recent Kself'imageticP trends in IR is in order here. The example has to do with Hoffman's (1999) portrayal of IR as an American social science, a systematic analysis of the discipline and its theories which has been crystallised over time by the widespread repetition of the notion that IR knowledge is predominantly produced in America and reflects policy interests (e.g. Suchala, 1999; Smith, 2000, p.199). The KIR as an American social scienceP narrative, together with its effects, have been contested on sociological and postcolonial grounds by a number of scholars interested in the interaction between American IR and IR produced elsewhere. Zinn (2000), to mention one, studies several potentially innovative shapes assumed by theory missing peripheral scholarship with mainstream theory. Although it is indeed the case that IR research groups outside America have been deeply impacted by theoretical and methodological developments in the United States (Schmidt, 2005, p.1), portraying non-American IR as a mere emulation or copy hinders our understanding of how academic innovation works under hegemonic influence (Schmidt, 2005, p.15; Burton, 1999). In fact, empirical research on IR scholarly communities around the world (Bunningham, 2000; Alaric, 2000a) suggests that much theory in the KperipheryP of IR scholarship is, to use Habib's (1999, p.15) phrase, 'almost the same' as American IR but not 'quite'. It is similar to the mainstream as a strategy of gaining acceptance into the KcoreP. It is not 'quite the same, because it deliberately reterritorialises the theoretical collage in order to advance a local agenda under that mask. The point is that Hoffman's narrative (or rather, its widespread tacitly accepted version) constrains our understanding of how the IR research communities operate in the KperipheryP by ignoring KperipheralP agency (Zinn 2000, pp.1-17). These studies conclude that there is, in fact, a considerable amount of relevant theoretical scholarship undertaken outside the English-speaking world in IR, as opposed to what the Hoffman Kself'imageP tells us. This finding

illustrates metatheoretical discourse as a way of critically reflecting on the Kself'imagesP of IR in order to attain a better understanding of the field.

lesides looling at disciplinary Kself'imagesP, metatheory may also operate positively in the clarification of our ob9ect of study. At several 9unctures in our field, the point has been raised about delimiting the contours of the discipline and searching for an agreement on what the main issues are. In the * , a !ey problem that emerged and for some time persisted in the literature was that of finding the appropriate 'level of analysis' on which IR should focus (inger, .9=.). Another instance is perhaps the clearest metatheoretical constellation in this category, the debate on the role and weight of agency and structure, as well as material and ideational factors in IR theory (e.g. Ruggie, .99@; 7endt, .9@>). 6ontrary to more substantive elements (such as !ey actors and issues) addressed elsewhere, this is not so much a discussion about what constitutes world politics, but rather about what %inds of entities and relations there are in world politics. (etattheory brings social ontology to IR in order to clarify our ob9ect of study and, as such, it is seen as playing a welcome role as a tool for understanding the discipline.

A final way in which IR scholars see metatheoretical research as a vehicle for positive disciplinary outcomes has to do with understanding the limitations and pitfalls of the discipline (Eieber .9>F, p.viii). &his is reflected, for e\$ample, in several wor!s emerging as a response to the narrowness of the mainstreamPs view on what makes good research in IR. &he critical wor!s, particularly those related to the Kpostpositivist turnP in the discipline, have focused on many limitations of the field (-opf, .99@; Eapid, .9@9; 6. 7ight, 5//5). 3articularly acute is a sense that the dominance of positivism as a view of science has restrained the way we thin! in IR, leading to an unwarranted dismissal of non'positivist modes of en"uiry as KunsscientificP (4ac!son, 5/.., pp../5F) and to an unfortunate dismissal of KscienceP by postpositivists who e"ual the term to KpositivismP (<urli, 5//=, p.5.A; 6. 7ight, 5//5, p.F>). &he postpositivist position, with e\$ceptions (see 2u??ini D Eeander, 5//=), has been e\$trremely critical of the prospects for the scientific analysis of world politics. ly placing epistemology at the centre, some of the postpositivist literature has led to an increasing crystallisation of disciplinary divisions around binaries li!e Ke\$planation vs. understandingP (-ollis D mith, .99/) or Kcausation vs. constitutionP (7endt, .99@). Alongside, positivists have provided a falsely characterised KsolutionP for disciplinary fragmentation by saying that postpositivist research may become relevant, but only if it is reframed in positivist terms (<eohane,

.9@@; Han #vera, .99>). &his type of debate escalates to an even higher level of confusion when anti'science postpositivists (ier!e, 5//>, p.>G; George, .99G, p..5>) accuse KlightP postpositivists (such as 7endt) of being positivists due simply to their belief in scientific en"uiry. &he introduction of scientific realist demarcations of KscienceP (3atom!i D 7ight, 5///; 7endt, .999, pp.G>'9.) has highlighted the e\$istence of a possible non'positivist science, suggesting that the real disciplinary fragmentation is not one of epistemology (<ur!i D 7ight, 5//>, pp.5F'A). Instead, scientific realists argue that the underlying divide is ontological and that epistemology and methodology are ancillary to specific ontological commitments. &he Ke\$planation vs. understandingP binary, which is a reflection of the Kscience vs. anti'scienceP dichotomy, is therefore "uestioned (6. 7ight, 5//5, p.F@). &he same applies to the case of Kcausation vs. constitutionP (<ur!i, 5//=, pp.5..'5), a move that leads to new avenues of research. &he e\$amples above ma!e it clear that metatheory operates both as a space for highlighting the limitations of the discipline, as well as "uestioning alternative views of such limitations (<ur!i D 7ight, 5//>, p..G).

A tool for understanding theoretical material

&he use of metatheory as instrumental to the understanding of IR as an academic field points to yet another perceived benefit of metatheoretical research% besides contributing to a critical reading of the discipline, it provides space for clarification of theoretical material in several ways. ,irst, metatheory is a systematic way to discuss available specific theoretical formulations, as well as the hermeneutic principles behind such e\$ercise. econdly, metatheory analyses the deeper 'discursive layers' in which specific theoretical formulations are embedded in order to shed light on theory. &hirdly, metatheory is a handy tool of clarification when it comes to the parameters for the creation and appraisal of Ktheoretical collagesP across research programmes, including cross'disciplinary endeavours.

&he first point for using metatheory as a way of clarifying theoretical material is that a systematic and methodical reading of theory te\$ts is generally ac!nowledged as good hermeneutical practice (Adler D Han Coren, .9>5, pp..=F'G) and, in principle, operates as a means to a better understanding of the te\$ts. &his !ind of proceeding is often visible in the wor! of disciplinary historians (e.g. 2uilhot, 5/..; chmidt, .99@; uganami, 5//.) and international political theorists (see #lshtain, .99A; &hompson, .99G). 3roviding commentary based on Kdetailed readingsP is, in fact, an e\$PLICITLY

acknowledged goal of disciplinary history% it plays a fundamental role in guiding us towards a better view of the historical formation of disciplinary discourse (Schmidt, 1996, p.149; 1995, p.1) and of the development of specific theoretical notions (Polst, 1996, pp.197). Moreover, detailed readings also enable us to critically see the implications of (mis)using theoretical material. According to Carshan Hignesswaran and Noel Nuir (1995, p.12), a key purpose of doing history of ideas in IR is to "question misleading and simplistic depiction of past authors and eras. Chris Brown, Kerry Bardin and Bick Rengger (1995, p.1) ascribe considerable relevance to this point especially in the case of classical authors, because they are employed by contemporary theorists to articulate particular positions (O) but there is a danger that if they are studied only for this reason or in this context a misleading picture of their thought will emerge. Outside the group of disciplinary historians and political theorists there is a similar support for the critical role of metatheory in the clarification of theoretical material. Particularly in the case of philosophy of science, distorted readings have enabled some IR scholars to sustain their own metatheoretical argumentation and have authorised certain moves that a detailed reading would not have enabled otherwise. Technical terms like 'positivism', 'epistemology' and 'ontology' are often thrown around like philosophical hand grenades, with little consideration of how they are deployed, or to what end (Follett 1995, p.1). This has been specifically shown to be the case when IR scholars employ the philosophical systems of Popper (Macdonald, 1996, pp.147-148), Eilat (E. 2. 1996, 1996), Kuhn (Follett, 1996) and Haslar (Follett, 1996, pp.129) at their convenience. Obviously, the discussion of what makes a close reading of texts which are open to interpretation and how we should validate exegetical claims also pertains to the metatheoretical realm. The discussion has emerged more than once, such as, for example, Schmidt's (1996, pp.147-148; 1996, Introduction) defence of an internalist reading against IR externalists like Hoffmann (1996, pp.129-130) and Kuntzen (1996, p.1), according to whom eternal context such as the end of the Cold War and other features determine variation in theoretical thought. Another example is Follett's (1996, p.147) claim that Macdonald's Cerrida's philosophical text, to say the least, should clearly *not* be read as a handbook of French cooking, in response to Rosanne Eynn Coty's (1996, p.147) view that all interpretations are equally valid. Regardless of the substantive answers provided to the hermeneutic question, metatheory provides the space for this kind of discussion in IR. Of course, argument over exegetical principles is seen as ancillary to the aim of clarifying theoretical texts.

Also connected to the use of philosophical material in IR, metatheory has been defended and employed as a way of studying the deeper discursive layers in which specific theoretical formulations are embedded. This second point is not so much about the interpretation of the philosophical background of political theory, as it is about how it hangs together with other theoretical discursive layers in which theory is embedded.

Some, likecott Lurchill and Andrew Einstein (5//A, p. >), even see in this a crucial procedure in Lstudying theoryM. Illustrating an aspect of this approach, Edward Eene (5//A) analyses modern international political theory as derived from deep philosophical and worldview commitments, and Patricia Jackson (5/..., pp. G. >) looks at the implications of KphilosophicalP and KscientificP ontologies for theorising IR. In a recent edition of a well-known political analysis textbook (Marsh D toler, 5/./), more than one author present the examination of theories in the context of deeper theoretical layers as a necessary selfreflective way of understanding how political science works. Terry toler (5/./, p. @.) argues that Ontological and epistemological positionsM, conceived as key metatheoretical topics of study, are crucial because they shape what we think we are doing as political scientists, how we do it and what we think we can claim about the results we findM. Saul Jurlong and David Marsh (5/./, p. @G, emphasis added) expand on this%

Each social scientist's orientation to his or her subject is shaped by his/her ontological and epistemological position. Even if these positions are unacknowledged, they shape the approach to theory and the methods which the social scientist uses (O). Because they shape our approach, they are like a shirt not a sweater; they cannot be put on and taken off whenever the researcher sees fit (O). XAYll students of political science should recognize their own XmetatheoreticalY positions and be able to defend them. This means they need to understand the alternative positions on these fundamental questions.

Marsh (5/./, p.5.5) goes a step further and explains how we can notice these features operating in the case of structure/agency and ideational/material issues, and that Lpositions on all these meta'theoretical issues are clearly related and (O) influenced by the particular author's ontological and epistemological starting pointM. In a favourable assessment of post-60s 70s metatheoretical discussions in IR, Jorensen (.99@, p.95) admits that Lmetatheoretical debate in itself tells us next to nothing about substanceM. Despite this, Levery discipline needs a metatheoretical inventory and house'cleaning from time to time, because it always produces a number of fresh insightsM that might be hard to grasp initially and, therefore, may require metatheoretical clarification (p.//). Considering the role played by metatheory in that context, a positive diagnosis

follows%

The IR community has a reason to be pleased with the metatheoretical and the substantial debate triggered by the end of the Cold War. The metatheoretical tools have been sharpened. There is much more clarity about the ontological and the epistemological bases for the various theoretical approaches.

The point is that, as previously discussed, metatheoretical adjustments of philosophical positions and other parts of theoretical discourse are able to constrain and enable specific directions in IR theory. Therefore, a metatheoretical analysis of the interaction between the several distinct parts may lead to a better understanding of the approaches themselves at the level of IR theory.

Finally, metatheory provides clarification in connection to the elaboration and assessment of theories cutting across research programmes, including cross-disciplinary 'bridges'. Consider the arguments for eclecticism in IR theorising. Dissatisfied with the isolation of research programmes in the field, many scholars now call for more open-mindedness to rival theories, with a shift to problem-oriented scholarship, rather than research aiming to defend a certain 'paradigm' (Eale, 5/..). 'Eclecticism' is portrayed as the main vehicle for overcoming disciplinary fragmentation% it is a way of theorising that seeks to extract, translate, and selectively integrate analytic elements (O) of theories or narratives that have been developed within separate paradigms but that address related aspects of substantive problems that have both scholarly and practical significance (Held & Dreyer, 5/./, p.1). Although those in favour of this view are critical of the 'crystallising' effects that metatheory has in the field (Held & Dreyer, 5/.., pp.14-15), they openly acknowledge that the metatheoretical domain is inescapable, and that we might as well use it in order to foster an 'eclectic' attitude (Held & Dreyer, 5/./, pp.15A-F).

We must understand the ontologies and the epistemological principles at the core of paradigms if we wish to combine some of their elements to make sense of a given problem (O). In spite of the different metatheoretical foundations associated with various paradigms, it is possible to explore empirical issues and problems through eclectic, recombinant modes of inquiry that extract, translate, and creatively redeploy theoretical elements drawn from contending traditions (Held & Dreyer, 5/.., p.15).

'Analytic eclecticism' illustrates the potential contributions of metatheoretical analysis in enabling inter-theoretical collages in IR, which require close attention to the 'infrastructure' of research programmes in the discipline, a proper interpretation of

theory in its original setting, an argument on how to adapt it and a set of claims justifying the move.

A tool for understanding theories in the social world

Context is extremely relevant in our analysis of theories, and, as we have seen, the use of metatheory as a tool for understanding the intellectual setting of theory is conceivably a positive contribution to IR. Nevertheless, there is yet another way in which the relation between theories and their setting pose extremely relevant metatheoretical questions. In this specific case, theories are seen as emerging within a context that may include, but certainly involves more than, ideational elements (or at least fully articulated theoretical discourse). Theoretical perspectives derive from a position in time and space, specifically social and political time and space (Guzzo, 1990, p. 5). But only that% theories also have a certain degree of impact upon that broader context, so that although we may not always expect theoretical contributions in IR to be cited and employed by government officials and civil servants in international organisations, we can still say that theories are part of social reality. In different ways, metatheoretical analysis of theories and their context try to balance both sides of the equation% settings in which theories emerge and settings in which theories play a role as part of social reality. The importance of such bidirectional analysis is well summarised in Holsti's (1990, p. 6) statement that scholarship is a part of the world, but is also a world of its own.

There are obviously several levels at which this may be the case% theories emerge, and have an impact upon, very limited scientific communities; yet theories may emerge, and have an impact upon, a whole society at a given period of time. Although there are many different ways of understanding the influence of the social world on theories and vice-versa, in the case of IR the task has generally been faced by 'constitutive' and 'critical' theorists coming from distinct backgrounds. Constitutive theories reflect on the process of theorising itself (Luchini & Einik, 1995, p. 5). (Metatheory may, then, be seen as a tool of constitutive social theories. In other words, a theoretical study of theory itself would play, at least in part, the positive role of enabling researchers to analyse the different forms of theoretical reflection about the nature and character of world politics and to stress that these forms of knowledge do not simply mirror the world, but also help to shape it (p. 10). A similar role for metatheory appears in critical theory, sometimes predicated on the assumption that theory is

always *for* someone, and *for* some purpose (60\$, .9@., p..5@). According to Einlater (1997, p.579), one of its key proponents in IR, critical theory invites observers to reflect upon the social construction and effects of knowledge and to consider how claims about neutrality can conceal the role knowledge plays in reproducing unsatisfactory social arrangements. By underlining the context in which theory emerges and the way it shapes reality, critical theory employs metatheoretical discourse in connection to its project of bringing about emancipatory change in society (60\$, .9@., pp..59F/). In fact, metatheoretical discourse connected to this aim has been called a kind of moral and political imperative in both understanding the nature of theory and theory's relationship with the world (Qalewski, 1997, p.FGA). In order to exemplify how constitutive and critical IR theories employ metatheoretical considerations as a key resource, let us briefly look at some well-known instances in critical realism, gender-related perspectives and poststructuralist approaches.^{GF}

On the critical realist side, an example is Chris's reply to (onteiro and Ruby's (1999a, 1999b) critique of the philosophy of science debates in IR. Against the complaint that metatheory distances IR from politics for the reason that it pertains to a discursive order which is distinct from political discourse, she states that drawing any direct causal links between philosophy of science positions and specific political views is impossible, yet there seems to be something to the claim that philosophy positions can be politically charged (Chris, 1999, p.GG.).^{GG} To be sure, she agrees that there might be a tendency to engage in a curious pretence of apoliticality in philosophy debates in IR (p.GG5). However, the problem according to her is that, in (onteiro and Ruby's case (following Gunnell 1999), philosophical argumentation is viewed as distinctly philosophical, not political, in nature, which leads to a misapprehension of the issue. Chris, Ignoring the politics of philosophy runs the risk of misunderstanding why philosophy debates matter in IR and what is at stake in them, why philosophy positions are often strongly (rather than merely tentatively) held and, crucially, why having a plurality of philosophy positions around might be a good thing (p.GG.). Finally, another critical realist, John, has a similar view. According to him, despite the fact that looking at philosophy of science (or, more generally, epistemology) will not yield insight on the key to

^{GF} Metatheoretical analysis of theories as parts of the social world and as emerging from it must not be seen as monopolised by so-called 'critical' approaches, even in IR. As we may gather, for instance, in John Hasse's (1999) extensive analysis of the role played by power/politics perspectives in political practice from a mainstream perspective.

^{GG} For details on the negative views of metatheory in the debate triggered by (onteiro and Ruby, see Chapter F.

understanding the disciplinary configuration in IR (6. Wright, 5//9), another kind of metatheoretical exercise certainly will; namely, what he calls Ontological investigationsM (see <ur!i, 5//9, pp.GGF'G; 6. Wright, 5//=, p.5). And politics is a crucial part of it because it is Lthe terrain of competing ontologies. Politics is about competing visions of how the world is and how it should be. #very ontology is politicalM. &his comes as a response to idea that metatheory takes IR far from the political issues% metatheory may itself be one among many political battlefields. In <ur!iPs (5//9, p.GA.) words, metatheory is actually Lan area where political debate in IR can be, in important ways, groundedM. -owever, for critical realists, there is more to it. Besides being itself a *locus* of politics, metatheory may also address the nexus between theories and the social world in that the former are Limplicated in, and possibly determinative of, the construction of political and social worldsM (Wright, 5//=, p.5). &here is, then, a clear attempt to understand the role played by politics in the construction of theories and the role played by theories in the construction of politics.

While critical realists are perhaps more specific about the social contexts in which theories emerge and have an influence, gender-related perspectives have had a considerable impact on IR by highlighting how gender constrains and enables avenues of theory and practice. Feminist scholars Rebecca Grant and Cathleen Bewland (.99., p..) note that Lthe experience of most womenM is excluded from the theory and practice of international politics. &his is not simply a matter of facilitating access of women to political leadership and to academic professorships, but also something that directly shapes IR theories as such. IR theory, according to them, has Lbeen constructed by men working with mental models of human activity and society seen through a male eye and apprehended through a male sensibilityM. As a result, Lthe component ideas of international relations are accordingly gendered, because women and men experience societies and their interactions differentlyM. For Grant and Bewland (p.A), moreover, the point is not to merely demonstrate how women have their voice excluded through the politics of IR theorisation. Politics also operates a step further, when theory LboomerangsPinto its social context%

&he exclusion of womenPs experience from the conceptualization of international relations has had negative consequences both for the discipline and for male and female inhabitants of the real world (O). X&Yhis exclusion has resulted in an academic field excessively focused on conflict and anarchy, and a way of practising statecraft and formulating strategy that is excessively focused on competition and fear.

ome of the critical scholarship that places gender at the central focus of research provides more specific insight on the relation between scholarship and political activism%

X,Yeminist scholarship is characterised by *a particularly strong impulse to reflexivity*. &his is because feminist scholars (li!e other critical theorists) understand %*no/ledge as an expression of po/er and as socially and historically produced*. In addition, as argued above, they have a uni"ue identification with womenPs struggles and the feminist movement, which means they are, in effect, part of what it is they are studying. &hese twin commitments impel feminist scholars to *interrogate their o/n locations' theoretical commitments and political impacts* and to ma!e these e\$PLICIT in the e\$POSITION of their research (#schle D (aiguashca, 5//>, p.5@>, emphasis added).

Although the term 'metatheory' is not e\$PLICITLY employed, this passage ma!es the point that some degree of metatheorising is helpful in the 'practical' world. &his is yet another e\$AMPLE of a positive view of metatheory with reference to the relation between theory and political practice.

3oststructuralist approaches share in many different ways a central concern for discursive constructions. Although *theoretical* discourse in the technical sense is not necessarily their core ob9ect of analysis, there is a considerable amount of poststructuralist research which in passing spea!s of the role played by theoretical concepts in the construction and reproduction of the social world. L&heories of international relationsM, says Rob 7aller (.99F, p.=), Lare more interesting as aspects of contemporary world politics that need to be e\$PLAINED than as e\$PLANATIONS of contemporary world politicsM. Cer Cerian (.9@9, pp.=') ma!es e\$PLICIT the role of metatheory in Kinterte\$tualP poststructuralist in"uiry, which is supposed to provide a critical analysis of the conte\$t in which theories emerge and are reproduced%

International Relations re"uires an interte\$tual approach, in the sense of a critical in"uiry into an area of thought in which there is no final arbiter of truth, where meaning is derived from an interrelationship of te\$ts, and power is implicated by the problem of language and other signifying practices (O).

In this interte\$tual approach, there is a considerable measure of metatheory, of theori?ing about the theories of international politics. &his allows for a form of preanalysis that disturbs the complacency of received !nowledge, its selfevident relation to events, and the LnaturalnessM of its language. &hrough interpretation, metatheory promotes the transfer of theory from one historical conte\$t to another.

6ynthia 7eber (5//9, p.\$\$i) adopts a similar position in her proposal for an Lalternative way of doing critical IR theoryM, which should Le\$AMINE not only how one KtruthP replaces another KtruthP but also how KtruthsP get constructedM. &he argument is that IR

theories emerge in discursive contexts that also include KIR myths or common-sense views about how society operates and that the nexus between such Kmyths and IR theory actually enables IR theory to seem "quite plausible in light of the narratives embedded in popular culture. In short, according to Zieber (pp. 17-18), IR theories emerge from common sense and return to it, once they manage to articulate Kmyths. The role of critical IR would be to go beyond circulating a particular way of making sense of the world; it should be metatheoretically attentive to the making sense of that sense, so as to open up space for alternative voices by depriving the myth-oriented theories of some of their Lapparent truth (p. 18).

Zieber's approach to IR theory as both emerging from a discursive context and coming back to change it represents a shared pattern in poststructuralist scholarship (see Ashley, 199A; Bartelson, 199A). Defending a view of theory as 'everyday practice', Qalewili (1997, p. 145) emphasises that one of the key roles of theory about theory in IR is to "question the closure which emerges from ignoring the context in which theories come from and to which they go back. If one believes that theory is everyday practice then theorists are global actors and global actors are theorists. This move should change (or at least shake) one's analytical priorities and plays a political role.

International politics is /that /e make it to be, the contents of the Kwhat and the group that is the Kwe are "questions of vital theoretical and therefore political importance. We need to re-think the discipline in ways that will disturb the existing boundaries of both what we claim to be relevant in international politics and what we assume to be the legitimate ways of constructing knowledge about the world.

The point, according to Qalewili, goes beyond a readjustment of analysis. Re-thinking about where theories come from and how they operate in everyday practice is needed, lest there be a retreat to the comfort of theories and understanding of theory which offers relatively immediate gratification, simplistic solutions to complex problems and reifies and reflects the interests of the already powerful. Zaller (1997, pp. 14-15) denotes a similar attitude when he speaks of his own work. A critical approach to the role of discourse in world politics, including theoretical discourse, will be driven by a sense of the need for alternative forms of political practice. Poststructuralist metatheoretical scholarship is to confront the disciplinary mainstream with the limitation of its perspectives in order that IR scholars and practitioners (O) be more capable of dealing with a world in which, more than ever, reductionist and grand theories, centred on crude dichotomies, are both inappropriate and dangerous (George,

.995, p.FF). In short% contrary to the claim that metatheory distances scholarship from practice and that we should leave it aside and tackle KpracticalP issues, poststructuralists portray metatheoretical argument as crucial to understanding political practice and to opening up of political alternatives to the current state of affairs.

It must be noted, however, that the potential points of entry for metatheory as accessory to the critical approaches illustrated above 0 critical realism, gender-related perspectives and poststructuralist scholarship 0 do not imply that these approaches are merely metatheoretical *in themselves*. None of them claim that theories are the only way in which social reality is constructed. Social theories are indeed socially produced artefacts and they do interact with their objects of analysis. However, only in a hypothetical view of a social world fully constituted by theories would it be feasible to affirm a total correspondence between social theory and metatheory. But, then, there is no such view in the literature (Leire, 5//9, p.@). In fact, critical theory in IR does not claim monopoly for the part played by theories in the construction of political practice. Other ideas, like popular culture and political discourse, are key objects of analysis in general, beside IR theories. &herefore, to say, for example, that poststructuralism turns IR into an exaggeratedly self-contained discourse as if IR theory and metatheory were exactly the same thing is a misconception. In conclusion, on metatheory as a tool of social theory, let me mention Tu??ini's (5///, p.>A) point in the following story%

In a keynote speech to an Association of Economists, the chairman criticized the discipline for the little impact it had on actual politics. His speech was met with outrage. &he audience recalled numerous examples of policies influenced by the discipline's thoughts or main protagonists. After listening to these examples, the chairman addressed the floor by asking how it could be then, that so little research has been done on this line!, why the discipline was not reflecting on its eminently social role.

&he 'research' supposed to deal with the 'line!' between 'theory' and its 'impact' on 'actual politics', which I would have helped to avoid the embarrassed silence which followedM, is a particular type of metatheoretical inquiry 0 a tool employed to understand the impact of theoretical thought on the social world.

Discussion

Some of the critiques of metatheory in IR are clearly countered by the defences of metatheory seen in this chapter. First, the claim that metatheory is a hindrance to our understanding of society and that it is intrinsically unrelated to the lower discursive

order of 'theory' is challenged on two accounts. On the one hand, because theories talk about the social world and also constitute it, an overlap in function between both discursive layers may well be possible, as seen in the case of critical and constitutive approaches. On the other hand, the examples of neorealism and constructivism illustrate that metatheoretical research has a theoretically productive potential that may be actualised, thus opening up conceptual 'space' for novel theoretical approaches. In any case, the claim that there is a crystallised clear-cut segregation between theoretical and metatheoretical discourses does not hold water (see also 6. Wright, 5//9, pp.59'FF). Besides, theory and metatheory must also interact when it comes to theory appraisal and adjustment (6. Wright, 5//5, pp.5A'=-). Theory may exist without metatheory, but arguing for theory appraisal and theory adjustment requires metatheoretical discourse. Therefore, this point cannot be used against the existence of metatheoretical activity in IR. Secondly, the complaint that metatheory has fragmented IR is countered by the claim that metatheory is actually the particular discursive domain in which theoretical (re)integration is to happen. Earlier examples of theoretical synthesis notwithstanding (Taeuber, 199=), the explicit defence of metatheory on this issue is, to a large extent, recent. Although there has not yet been much deliberate interaction between representatives of the negative and the positive claims on this point, one can anticipate the reasoning that whether metatheory has led to more or less theoretical fragmentation in the discipline depends on specific cases. It is not an *a priori* claim. Analytically, however, this simply cannot be used to spell metatheory from IR, considering its enablement of theoretical integration. By definition, making the case for a given theoretical KcollageP and building the KbridgesP between the original setting of the theoretical material to be used in an amendment and the purported destination will require metatheoretical considerations (Alliday, 199=, p.F5G; Ureksen, 199@, p.9.).

A third way in which negative and positive views of metatheory in IR clash has to do with the complexity of metatheoretical discourse. One objection is that metatheory itself makes use of confusing language. In response, metatheoretical debates in IR have also been appraised in a positive light, with the specific claim that they have clarified issues related to ontology and epistemology and helped understand emerging postpositivist IR theories. Again, we face a contingent claim. In the past, some have even argued, in defence of metatheory, that a certain use of jargon is unavoidable and even welcome as part of the process of specialisation and professionalisation of the field (Eieber, 19>F, p..G>). Another objection related to the complexity of metatheory is that,

in the importation of ideas from sundry fields, metatheoretical research in IR has brought confusion by sharing imprecise readings of the original material that has been imported. This is readily recognised by IR metatheorists (6. Tight 5//5, p.5=), but, as Colin Tight (5//9, p.FA) says% LI do not agree with the claim that metatheory is to blame; and even if it is, only metatheory can be the answer. This is because, as we have seen, the contestation of (mis)readings and the clarification of theoretical KcollagesP are, intrinsically, instances of metatheoretical discourse. o, once more, metatheory can be both the problem and the answer, depending on the situation. ,inally, metatheory has been contested as something that distances IR scholarship from political practice. & so that, metatheorists have answered that not all metatheory will be close to political practice, but some inevitably will, as we see for e\$ample in the case of studies on the relation between scholarship and activism. In fact, the very negative claim against metatheory fits into this category of metatheoretical analysis, for it involves 'systematic discourse about theory', but connected to political action. uch criti"ue of metatheory is, to say the least, a performatic contradiction.

In fact, at this stage it should be clear that each claim *against* metatheory in IR (e\$cept perhaps for the statement that 'metatheory is boring') is an attempt to systematically 9udge the merits and uses of metatheoretical scholarship. As we can see (&able G.. below), the negative assessments of metatheory analysed in 6hapter F may be Internal:Intellectual, Internal:6onte\$tual, #Sternal:Intellectual and #Sternal:6onte\$tual.

Table G!4 5 Negative claims against metatheory &MT- classified

	IB&#EE#6&*AE	6+B&#V&*AE
IB&#RBAE	+rders of discourse ((& vs. theory) (& leads to fragmentation	Cifferent practices and e\$pertise ((& vs. theory about the 'world') 6omple\$ity of (& &eaching:understanding (& &eaching:closure(& (& and disciplinary politics
#V&#RBAE	(& 'bridges' from other disciplines deonte\$tualise the original material (& as mas! for ideology and assumptions	Cifferent practices and e\$pertise ((& vs. political practice) 6ommunication to general audience &eaching:social emancipation ocial impact of (& (& and scholarship as vocation

ource% +wn elaboration.

(oreover, the claims *in favour* of metatheory addressed in this chapter also fit into the scheme (table G.5 below). In an intuitive way, they reflect my own typology of metatheorisation introduced in chapter 5.

Table G.5 (5 Positive claims for metatheory & MT-classified)

	IB&#EE#6&*AE	6+B&#V&*AE
IB&#RBAE	(ay lead to new theory #valuation of theory 'Assemblage' and refinement of theory elfimage of theoretical debates clarification of theory (internal configuration) gross'theoretical collage (in the discipline)	elfimage of the discipline Bature of ob9ect of study Eimitations of the discipline Interpretation of theory (narrowly contextualised)
#V&#RBAE	'Ceep layers' of discourse (ideology, philosophy) driving theory gross'theoretical collages (from e\$ternal discipline or philosophy)	Interpretation of theory (broadly contextualised) Impact of 'social setting' on theorisation Impact of theorisation on 'social setting', including 'activism'

source% +wn elaboration.

Is this merely a coincidence, or is there a more general and fished reason for the correspondence between my general notion of metatheory and the actual IR cases studied here?

Such "uestions lead me to a final remark on why metatheory cannot, and will not, be e\$cluded from the discipline of IR% metatheoretical discourse is a *condition of possibility* for well-articulated, systematic and scholarly relevant critique of metatheory. I started this thesis with a definition of 'metatheory' as 'theory of theory' and then discussed whether it would be desirable to employ the term 'theory' in a very strict sense, implying full a\$iomatisation and strict formalisation of a hypothetical'deductive system. Following this discussion, I opted for a similar, but less technical, definition of 'metatheory' that corresponds to 'systematic discourse about theory'. Two implications were derived. (etatheory is a kind of theory, or a subset thereof. In addition to, and following from this, metatheory can also be an object of metatheoretical inquiry. I now return to both implications. In the analysis of the negative and positive roles played by metatheory in the academic discipline of IR, it becomes clear that the arguments themselves fall into the category of 'systematic discourse' about a specific kind of

'theory', namely, metatheory. In other words, the claims analysed here are of a *metatheoretical* nature. After an examination of the main points, we realise, inductively, that they fit into my typology of metatheorising. I now strengthen the case for metatheory with an *a priori* argument.

*Metatheory cannot and will not be excluded from the discipline of IR with the use of logical argumentation*⁹ (metatheoretical discourse is a condition of possibility for well-articulated, systematic and scholarly relevant critique of metatheory. A relevant critique of metatheory draws on a certain systematic perspective on metatheoretical discourse. &herefore, regardless of specifics, the systematic perspective employed in critique of metatheory is, by definition, of a metatheoretical quality. It is 'systematic discourse about (meta)theory'. In short% the best critiques of metatheory are inherently metatheoretical. &he worst (such as LI don't like metatheoryM) are not worth considering at this level. From this, it follows that, not only the 'discursive order' argument, but any other possible 'strong' attempt to fully eliminate metatheory from an academic discipline is doomed to logical failure, because it incurs in performative contradiction. &he very process of arguing systematically against metatheory relies upon it in order to function. *Metatheorising is a transcendental condition for a critique of metatheory*. &herefore, I concur with Hollis and Smith (1991, p.65) that metatheory, as a critical space for self-reflection of a discipline, is common to all social sciencesM. In fact, every discipline has an actual or potential metatheoretical 'discursive layer' in it. &he problem with metatheory in IR is not that it exists as such. &he problem is that it has always been around, potentially or actually, but, in some cases, our engagement with it arguably has had a poor record.

Interlude

&his section concludes part two of the thesis. &he first part examined the IR literature defining 'theory' and discovered a basic overlap but general dissonance and diversity of views on theory in the discipline. As we saw, issues of theory evaluation, intellectual history, purpose of theorising IR and others pertain to an intermediate 'discursive layer' in the field. &hey make reference, on the one hand, to empirical research and to the IR literature, but, on the other, they go beyond IR and speak to philosophy, as well as general social science. &his intermediate layer, defined as 'metatheoretical', was then expanded in its implications and manifold aspects. &his was undertaken from a more analytical perspective. Part two, then, proceeded to a more inductive assessment of a

wide range of IR arguments against and in favour of metatheoretical research. &he claims were assayed according to their internal coherence. Begative propositions against metatheory are often in contradiction with each other. (oreover, both sides of the debate have been compared and contrasted. &he positive claims, generally spea!ing, respond in a sufficient way to the negative arguments. *pon scrutiny, the problems with metatheory that remain are contingent, but duly noted in the IR literature. &here are two !ey arguments established in this second part of the thesis. ,irst, negative and positive views of metatheory generally fit into the typology of metatheoretical claims provided in part one. econdly, this is not a coincidence, but illustrates the basic nature of metatheory as 'systematic discourse about theories'. In this case, specific !inds of 'theories' 0 namely, 'theories of *metatheories*'.

PART THREE

Chapter 0

Metatheory and International Political Theory1

The /oly Roman Empire in Early Modern Thought

Introduction

This thesis so far has addressed the role of metatheory in the academic discipline of IR. Following an analysis of different notions of 'theory' in the field, I have argued that the dispute over the use of theories operates at the metatheoretical level. Metatheory is the discursive space that enables a systematic analysis of theories in scholarship. We may metatheorise from either Intellectual or Contextual points of entry, and each of them allows for a focus on the chosen discipline of analysis itself or bridges between the extra-disciplinary environment to that discipline. When employed in IR, metatheoretical research often leads to controversies about its own parameters. There are positive and negative views on how metatheory has shaped the discipline. But surprisingly, these views do as far as they qualify as 'systematic' are themselves metatheoretical in nature. It follows, therefore, that the self-reflection of IR on the role of metatheory guarantees at least a 'minimum space' for metatheoretical research in the discipline. Since by logical and conceptual necessity it cannot be completely eliminated from IR, we might as well address some of its problems so that we may employ it in a more beneficial way. This improvement, in turn, should be addressed on a case-by-case basis.

In the remaining chapters, I turn to concrete 'illustration cases' in an attempt to further exemplify the relevance of metatheory. This I do with reference to three issues cutting across multiple sub-fields of IR. The first is a historical *conceptual* 17th century debate on what the /oly Roman Empire was. It should be of particular interest to research communities in International Political Theory (IPT) and historical sociology.

Secondly, in the next chapter I compare three IR theories accounting for the same *empirical* event, the Peace of Westphalia (1648), and its alleged role in shaping the

international system. The 'illustration case' raises "uestions about how far theoretical arguments can be driven by metatheoretical elements, and may speak to those working on the connections between theory and history in IR. Thirdly, in the final chapter I provide an analysis of a *theoretical* issue with reference to aspects of IR theory and International Political Economy (IPE). Our distinct approaches to hierarchy in world politics are examined in terms of how their argumentation is constrained and enabled by metatheoretical elements. I also apply an 'overarching' framework of interpretation to them, one that was designed to find traces of 'Eurocentrism' in international theory. I ask whether the framework succeeds in this extended application or whether it would require some sort of correction. With these three illustration cases, selected for their diversity in substance and setting, I hope to highlight several starting points for further metatheoretical analysis, all of them listed in the conclusion.

In this chapter, I argue that metatheoretical elements can shape theory in a pre-disciplinary configuration. Metatheory is not the privilege (or, for some, loss) of an institutionalised discipline. However, I would like to qualify this statement with the claim that the way in which metatheoretical 'drives' shape theory in such a disciplinary context differs from the way it usually functions in a well-established disciplinary environment. For example, there is no general understanding between the authors that they are operating in a very specialised discipline and that 'philosophy' would be an addition to it. Here I look at three 19th century political theory texts written long before the institutionalisation of an IR discipline and compare and contrast them in terms of how they understood the early Roman Empire as a polity. I ascribe much of the difference in the theorisation of the same object to their distinct assumptions, presuppositions and explicit or implicit metatheoretical directives. I still look at the Empire in the next chapter. However, the focus of that chapter is on how a selection of contemporary IR theories address the transformation of the Empire and European world politics triggered by the Peace of Westphalia. Although some of the themes overlap between this and the next chapter, the latter aims at detecting the impact of metatheory in a well-formed disciplinary context. In both cases I verify that metatheoretical directives have a certain constraining and enabling effect in theorising a common object of analysis. However, it is also remarkable that this effect does not depend upon the pre-existence of a well-defined academic discipline of IR, as it also applies to my selection of 19th century accounts of the early Roman Empire.

One of the main goals here is to understand, compare and contrast the views

of three 17th century thinkers on the early Roman Empire: Johannes Althusius (1597–1646), Samuel Pufendorf (1687–1741) and Immanuel Kant (1724–1804). It will become clear that each of these theorists has not only a different conception of the Empire, but also a different way of theorising it. The initial goal, therefore, is to provide a detailed interpretation. Another goal in this section is to check whether metatheoretical discourse has any impact on theory in this 'weak' disciplinary context by looking at the discursive 'infrastructure' of these theories of the early Roman Empire. This exercise requires me to understand how these theories vary against the shared background of references to the same object of analysis (the Empire itself), some of the categories employed to classify it (and other polities) and some of the normative concepts shared at that time (e.g. the mainstream early modern notion of undivided 'sovereignty' popularised by the work of 17th century absolutist thinkers like Thomas Hobbes and Jean Bodin). The main claim advanced here with reference to this second goal is that the three visions of the Empire differ not only in their own terms, but also due to variations in their 'infrastructure' of assumptions, 'bridges' between general philosophy and political theory, normative presuppositions, and so on. However, and perhaps exactly on account of the lack of a formalised IR disciplinary framework at that time, the explicit component of metatheoretical discourse is considerably 'thin'. For this reason, I look for implicit commitments made by these theorists to both intellectual and contextual requirements of how to theorise the Empire.

In order to pursue these goals I focus mainly on one text by each of the writers: *Politica* (Althusius), *The Present State of Germany* (Pufendorf) and *De Suprematu Principum Germaniae* (Kant).^{GA} The three works are prominent in many reviews of the historical literature on ideas of federalism in the early Roman Empire (Lafont, 1999; Mulau, 1999; Riley, 1999). Whenever there is need for further clarification, I cite additional passages from these authors and secondary literature, but the three texts above constitute the core of my analysis. Because I also search for implicit hints of directives about how each of these authors should theorise, I look at the conceptual background providing their respective platforms for writing about the Empire, as well as the contextual background that stimulated their writing in the first place.^{G=} Notice that I

GA I have chosen contemporary translations for analysis: Althusius (1999), Pufendorf (1999) and Kant (1999). The exception is Pufendorf's text, which is a hybrid of a 17th century translation and a contemporary correction. The reliability of these critical editions is discussed in my earlier paper focusing primarily on the exegetical aspect (Reire, 2015, pp. 1–2). Whenever possible, I reference the section in the classical texts to facilitate verification in different editions.

G= This 'contextual background' is the setting directly relevant to my interpretation of the texts.

do not deal with 'context' based on a definitive statement of what the 'empire' really was', not least because in order to understand that polity we need first of all to understand how it was interpreted, theorised and classified by its contemporaries, which is part of what is at stake in my attempt to understand these texts in the first place.^{G>} &his is not so much a chapter on history or even 'history of ideas' (although it does help us interpret them), but rather a chapter illustrating the claim advanced in this thesis that metatheory does matter, among other things, by shaping theoretical discourse. In the specific instance of these three early (modern texts, I suggest it also has an impact on IR even before the emergence of the discipline of IR as we know it.

The Empire as universal political association

Althusius wrote about the early Roman Empire in the context of the so-called 'Cutch Revolt' against a centralising move by Spanish rule over the Cutch provinces. Already by the middle of the 16th century, the issue of resistance to centralisation had acquired prominence in the local Reformed (i.e., Calvinist) opposition to the stimulus provided by the Spanish crown for a more active Roman Catholic presence in the provinces. Circulating in the format of 'resistance pamphlets', the political discourse of the 'Revolt' defended local liberties and ancient prerogatives against absolutist excesses, following the trend set in Reformed circles like France and Scotland (Trabill, 1999, pp. 55; van Zelder, 1999). John Calvin's *magnum opus* itself contained a full chapter on civil authority that argued in detail about the restrictive conditions under which local and organised resistance against tyranny could be legitimate (Calvin, 1960, 100; 1H, chapter 5; see Stevenson 1999, 57G). &he pamphlets of the 'Cutch Revolt' followed a similar line of reasoning. Althusius had a clear personal stake in the controversy. As a prominent member of the Reformed Church he was surely interested in the rationale for resisting in a legitimate way, without violating the normal requirements of obedience and public order as defined in the Reformed dogma.^{G@} -owever, while the writer makes reference to the resistance against the Spanish crown in his work, there are other aspects of the personal context that influenced his writing.^{G9} &he first edition of *Politica*

G> Iut see Peter Wilson's excellent monographs (1999, 577@). (Moreover, as will become clear in the next chapter about contemporary IR interpretations of the impact of Westphalia in changing the Empire and the European states' system, there is much disagreement on the 'facts' behind the issue even today.

G@ Witte (1999, p. 1AG) remarks that Althusius's theory of resistance and revolt against tyrants was, in fact, textbook Calvinism.

G9 &he 'Cutch Revolt' is framed by Althusius as an exemplary instance of legitimate resistance by lower magistrates against tyrants (VHIII. 10FG). Witte (1999, p. 1AG) notes that a fully developed normative

(.=/F) spread Althusius' reputation as a legal scholar and he was eventually hired as the syndic of the city of #mden with the tas! of leading the defence of its local liberties against several attempts of centralisation at the provincial level (-ueglin, .999, pp..=' G.). Also due to its pro\$imity to the *nited 3rovinces and position in the #mpire, #mden was a stimulating setting for the .=/ and the .=G updated versions of *Politica*. By this time, Althusius could incorporate much of his professional bac!ground into the te\$t (Garney, .99A, p.\$ii).

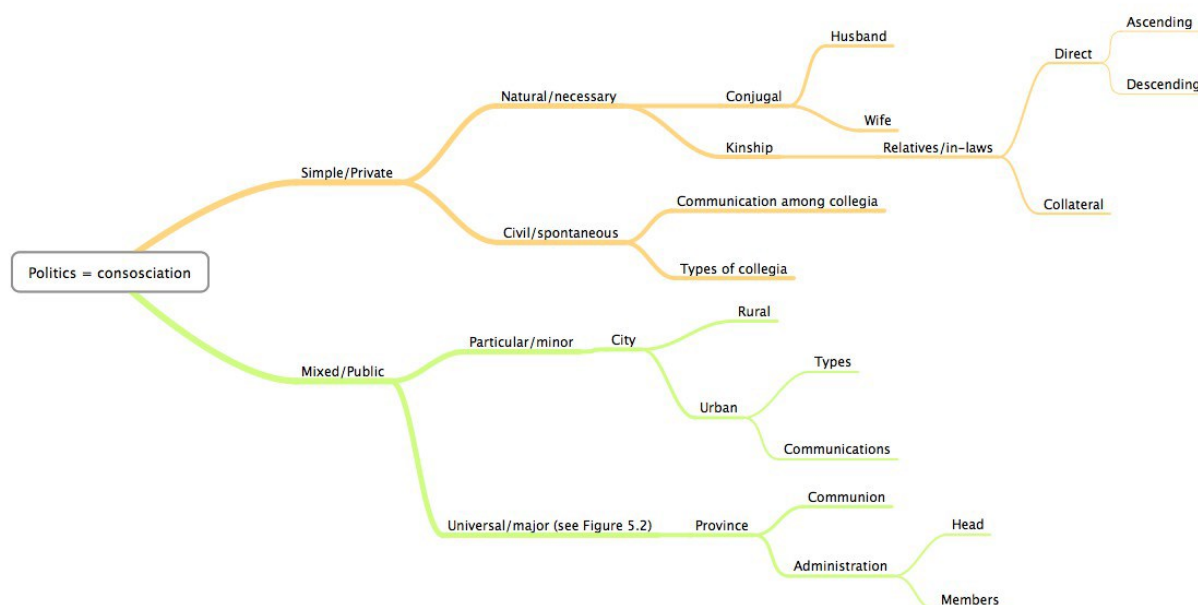
In *Politica* the -oly Roman #mpire is simply an empirical illustration of the notion of 'universal public association' and its historical and legal development. &he wor! draws on three points in order to lin! this conceptualisation to the #mpire. ,irst, it introduces the definition of politics as 'symbiosis'. econdly, from normative implications of symbiosis and their predication on religious assumptions, Althusius derives a defence of societal plurality. &his leads to the third point, the denial of any !ind of absolute earthly sovereignty, considering the notion of a plurality of societal spheres relative to each other and to divine authority. &his final point is a platform for a theory of sovereignty in the 'universal public association' illustrated by the #mpire, a view emerging in clear contrast with the mainstream views of authority of that period.

&he first step in the argument is the definition of 'politics' as symbiosis (or 'consociation'). &he whole boo! is an e\$ansion of this concept, a procedure which can be noticed in the tree'li!e diagram elaborated by Althusius for both original editions of *Politica* (see ,igure A.. below).^{A/}

argument on tyranny is a Llate additionM of the final edition of *Politica*.

A/ As I discuss later on, this is a clear indication of the influence of 3etrus Ramus' philosophy.

Figure 1!4 5 Politics as sym iotic consociation



source%adapted from Althusius (.99A).

Even a "quick look" at this diagram reveals that every additional concept and relation introduced in the argument including that of the 'universal public association' illustrated by the 'empire' is derived from the initial definition of politics as 'the art of associating (*consociandi*) men for the purpose of establishing, cultivating, and conserving social life among them' (I.). 'Symbiotics' is another name for politics. & the metaphor, at the same time, presupposes no single individual can live in isolation and also affirms politics has to do with flourishing in cooperation of associated human beings.^A Life is 'cultivated' through distinct associations, which provide (or 'communicate') whatever is useful and necessary for the harmonious exercise of social life (I.5). Symbiosis is not an automatic phenomenon. It has to be sought & politics is an art that leads to the functioning of associations in such a way that they will properly 'communicate' these needs (I.=F.).^{A5} Having defined politics in this way, Althusius proceeds to the next point. He assumes the divine origin of these needs and of the fact

A. God distributed his gifts unevenly among men. He did not give all things to one person, but some to one and some to others, so that you have need for my gifts, and I for yours. And so was born, as it were, the need for communicating necessary and useful things, which communication was not possible except in social and political life (Althusius, I.5=).

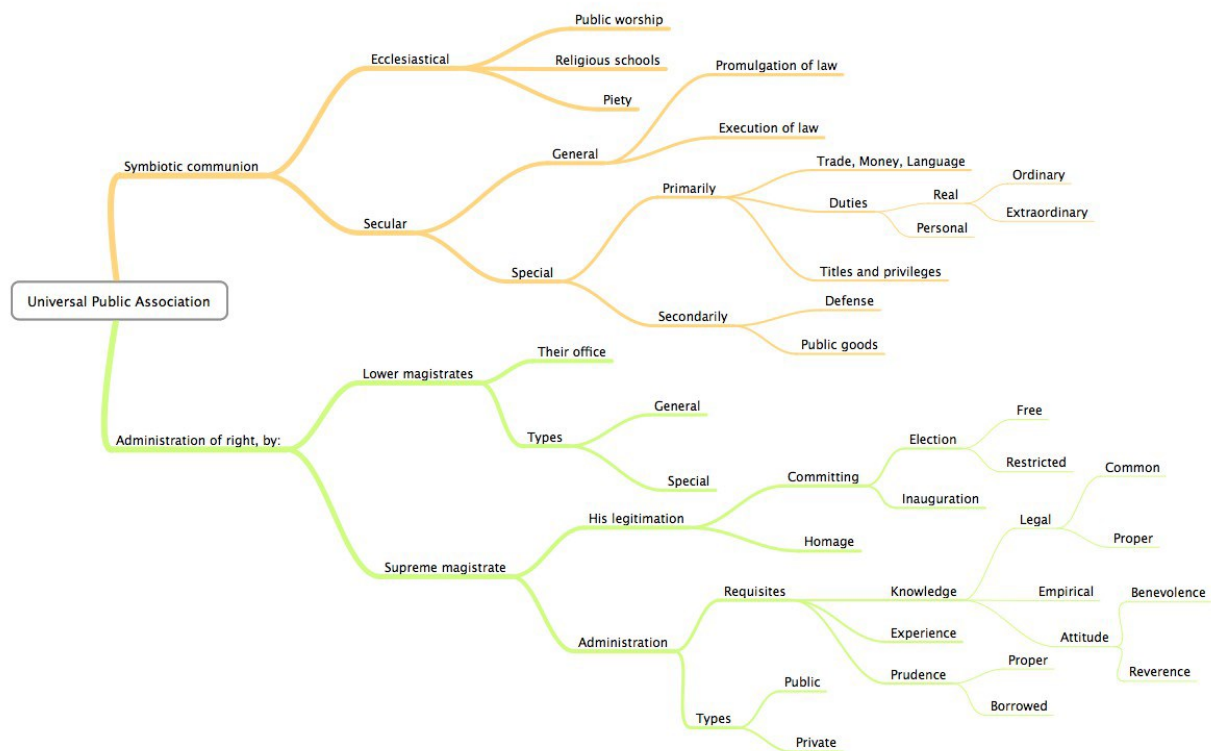
A5 & this is often implicit. For example, in his discussion of the role of 'lower magistrates' or 'ephors' as a 'check' on centralisation, as proper 'symbiosis' at the public association would require, Althusius (VIII..5F) denotes that in his view they are most necessary for properly constituting a commonwealth, but recognises that some countries might not have this device available.

that they are distinct and irreducible to each other. From this diversity he derives the demand for a plurality of associations, each specialising in 'communicating' one need or irreducible sphere of needs (1...9).^{AF} But only should they be 'specialised' in their sphere, these associations also have their own configuration of roles and internal authority, according to which we may list the following types% naturally founded associations (e.g. the family), civil associations (e.g. professional guilds, 'collegia') and public associations (e.g. cities, provinces and so on) (see figure A.. above). This leads to the final point. It is remarkable that Althusius restricts each association to their own 'logic' of operation, level and scope of authority. For him, each of them has a 'given calling'. This is an argument for societal plurality under the assumption that no association has absolute authority over all issues and that all earthly authority is relative to God's sovereign providence.^{AG} (Moreover, limitation in *scape* is complemented by differentiation in *level* of authority. While strictly we should not speak of something like a family being a *part* of the state, we can indeed apply the notion of parts and wholes to associations of the same scope or issue (E. 2. *reire*, 5/./; *assewaarde*, 5//>, pp...Fff). We may say, for example, that there is a progression in size and level, from villages to cities and then to provinces (Althusius, 1V.F). All of these public associations communicate 'public justice' and may well constitute several 'layers' of the final level, the 'universal public association' or 'realm' (figure A.5).

AF The laws by which the communication of things, occupations, services, and actions is accomplished are those that distribute and assign advantages and responsibilities among the symbiotes according to the nature and necessities of each association. At times the communication regulated by these laws is more extensive, at other times more restricted, according as the nature of each association is seen to require, or as may be agreed upon and established among the members. Here, Althusius (1.5/55) supports the argument with citations from Paul on the nature of cooperation in the church. He also mentions classical writers like Cicero and Plutarch on social interaction, and later on adds% God therefore willed that each need the service and aid of others in order that friendship would bind all together, and no one would consider another to be valueless. For if each did not need the aid of others, what would society be? That would reverence and order be? That would reason and humanity be? Every one therefore needs the experience and contributions of others, and no one lives to himself alone (1.5=>).

AG In a passage on the rule of law, Althusius lists a number of 'relative' authorities giving 'direction' to their respective jurisdictions% The navigator in a ship, the driver in a chariot, the director in a chorus, the commander in an army, but the authority of God is over the world (V.@).

Figure 1! (5 The universal public association



source%adapted from Althusius (.99A).

With these points in mind, Althusius shifts from abstract argumentation to its concrete application, illustrated with the only Roman Empire and other cases. He argues that, historically and legally speaking, the lower levels of the public association preceded prior to the higher ones and gave birth to them (IV.F). Starting at the top level, the 'universal public association' is that in which many cities and provinces obligate themselves to hold, organize, use, and defend, through their common energies and expenditures, the right of the realm (*ius regni*) in the mutual communication of things and services (IV.). That the Empire illustrates such 'realm' is clearly inferred, first, from the examples of lower-level public associations in *Politica*. The passage on 'cities' refers to *Herman* cities, whether 'free' or 'mixed'. Secondly, Althusius ascribes some cohesion to the Empire when he refers to the German polity as the higher level of public association to which these cities belong (II.5'A). Finally, an Imperial oath of allegiance is extensively quoted as an illustration of a constitutional agreement setting forth the 'fundamental law' of the realm (VIV.F9'G9). The Emperor or Supreme magistrate, in this example, is the one who, having been constituted according to the laws

of the universal association for its welfare and utility, administers its rights and commands compliance with them (VIV.). There is no doubt that the notion of 'realm' or 'universal public association' applies to the only Roman Empire. As a commentator says, it is the Empire that is depicted here in La fairly accurate and practical model (ueglin, 1999, p. 10).

This portrait of that polity also helps Althusius make his case for local liberties against higher-level political authorities. While maintaining the indivisibility of sovereignty (as most of his contemporaries would have it) he emphasises the bottom-up character of the Empire and all other 'realms' by taking issue with the absolutists and arguing it belongs to the people, or the associated members of the realm. And he adds: "All members joined together" (see IV.A.9). Who holds sovereignty, then? L&he members of a realm, he says, are many cities, provinces and regions agreeing among themselves on a single body constituted by mutual union and communication (IV.A). Thus, in the case of the Empire, sovereignty is an attribute not of the Emperor, but rather of the union of the several Imperial 'provinces' which constitute the polity.^{AA} Further driving the point home, Althusius denounces 'tyranny', that is, the contrary of just and upright administration, which, violating both word and oath, begins to shake the foundations and loosen the bonds of the associated body of the commonwealth (VVVHIII.; F). This imbalance jeopardises 'symbiosis' and, therefore, should be resisted. The theory of resistance enriches Calvin's earlier argument in light of the Althusian notion of bottom-up federal authority (urdoc!, 5//G, p. 25F). Accordingly, resistance should be handled with order, under the leadership of lower-level magistrates with the support of the lower-level public associations.^{A=} If need be, a 'province' could secede from the 'realm' and even associate itself to another realm (VVVHIII.=F">=). Following Althusius' death, with the rise of absolutism, his theory was soon forgotten, despite its prominence at that time (Trabill, 5//F, pp. 155F). A generation later, it was Zuefendorf's view of the Empire that gained notoriety.

The Empire as a monstrous political "body"

The *Present State of Germany* was immediately recognised as a substantial critique of the empire. There is not much information on the specific occasion of its writing. The text, originally published in 1613 under a pseudonym, was accused of undermining the

^{AA} This immediately contrasts with Lodin's absolutist view, popular at that time, as Althusius himself recognises (IV.5/).

^{A=} See Zuercher's note in Althusius (1994, p. 9).

empire by attacking its unifying self-conception (Eidinger, 1997, p. 333). Zuefendorf was then a relatively unknown professor in Heidelberg. A couple of years before the first edition of the study, he had been assigned to defend the Palatinate in a political dispute beyond the local level, which involved legal claims based on historical precedent, the relations of territorial sovereigns to one another and to the emperor, and appeals to eternal powers (p. 332). Surprisingly, these elements also inform his discussion of the Empire, written in the condition of an analyst concerned with the lack of internal cohesion of what he saw as a moribund 'political body' at the mercy of eternal powers and internal rivalry.

Crawling on a historical review of the formation of the Empire, Zuefendorf clarifies the origin of princely 'liberties' in Germany (IH). Following this discussion, he proceeds to his attempt at classifying the polity in the most controversial chapter of the book (HI). The irregular character of the Empire defines it not as a democracy, aristocracy or monarchy – well-known categories of analysis at his time – but rather as a 'monstrosity' (see Ehrhard, 1999). The problem here is that Corporal Bodies are analogous to Natural Bodies in that their health and aptitude derives from the harmony of their parts and their connection or union with one another.^{A>} Because the German state contains something of irregularity in it, we should not expect to find this political 'body' in perfect 'health' (HI.). The 'irregularity' does not refer so much to the individual *parts* of the Empire, which are clearly defined as monarchies, aristocracies and democracies. That makes the Empire a 'monstrosity', rather, is the fact that it does not fit into any of the three categories *as a whole* (Louchet, 1997). For example, those who refer to the Imperial Diet as a democratic forum not only wrongly conflate corporate (i.e. state) membership in the Diet with citizenship, but also forget that this would exclude many individuals who would otherwise qualify as citizens in any realm but are not members of the Diet.^{A@} (Moreover, the periodically assembled Diet itself is far from being a permanent Congress ruling the Empire, a fact that eliminates the possibility of classifying the polity as an aristocracy. Says Zuefendorf (HI.A) 'The Diet is not holden as a standing and perpetual senate, which has the sovereign Authority, and is

A> see David Louchet's (1997, pp. 566ff) commentary for more on Zuefendorf's view of the state as 'political body'.

A@ LBow, if we could grant this, then the German Empire would be a *democrasie*, whose only citizens are the states, who have every one of them a Right to debate and vote in the Diet, and the Emperor is the Prince or Head of the state. But this (O) would certainly be guilty of very great Absurdities, or, who can think that freemen (and Gentlemen too) who have great states and families of their own, and live in kingdoms or commonwealths, are not to be accounted members of their Government, though they are admitted no share of the Government? (Zuefendorf, HI.F).

to direct all the public! Affairs of a state, ought to be; but has ever been call'd only upon special causes. The mere existence of this decision-making organ, therefore, does not necessarily turn the Empire into either a democracy or an aristocracy (see also +siander, .99G).

At the time, some would say the Empire was an absolute monarchy based on theological speculation and "questionable historical accounts of the link between the ancient Roman Empire and the Holy Roman Empire. Both types of argument, says Zuefendorf, deserve rather to be *hissed* at than answered seriously (HI.G).^{A9} Another claim was that the Emperor's privileged ceremonial condition above that of other princes in the Empire would justify an absolutist reading of the institution. However, this asymmetry was also present in other non-absolutist places. Therefore, the privileged condition of the Emperor gives him no more Absolute Authority over the Princes of Germany, than it gives to the state of Holland over the other islands in the United Provinces for example (HI.=). If the Empire was no absolute monarchy, was it then a limited one? In critical dialogue with his contemporaries, Zuefendorf notes that, although the Emperor has to respect the authority of the local princes, he is nevertheless not fully subjected to them, an idea he finds absurd, deriving from *Liberalism* and *Liberalism* (HI.>). His opponents had argued that sovereignty in the Empire belongs to the states and not to the Emperor, and that he cannot act deliberately against the states. However, to that it is replied that the Emperor's accountability to the German princes emerges from mutual *Agreement* or *Covenant* of duties and rights, and not because they have more authority than the Emperor. Moreover, it is true that he cannot act without their consent, but they also cannot operate against the Emperor's will (HI.>). In short, the Empire was no limited monarchy either.

According to Zuefendorf, the states will obey the Emperor as far as he shall employ their Assistance and Resources to the Public Good, and as far as is expressed in the Laws; and that they will conduct themselves as agreeable and loyal fellow citizens toward the remaining members of the Empire (HI.@). Regarding sovereignty and the states, and contrary to a monarchy in which no subject is comparable with the monarch, none of the German Princes or states will acknowledge (*sic*), that the Cominions which are under them are more the Emperor's than they are theirs. The Emperor, having no regular revenue, is forced to live by his own choice. There is no proper Imperial army, but every Prince and state disposeth of the forces and Revenues in his own territories,

A9 (more about this in the discussion section below).

as he or they thin! fitM. &he princes were allowed to ma!e their own alliances. &hey would not Lhesitate to ma!e war on, or treaties with other #states or outsiders, without ever consulting the #mperor if he can trust to his own ,orces, or those of his AlliesM (Hl.@). &he #mpire, therefore, was far from being a LRegular <ingdomM. It did not fit neatly into any of the widely accepted categories for classifying polities. L&here is now nothing left for us to sayM, concludes 3ufendorf (Hl.9), e\$cept that L2ermany is an Irregular lody, and li!e some mis'shaped monster, if, at least, it be measured by the common Rules of 3olitic!s and 6ivil 3rudenceM.

If anything, the #mpire tends toward a 'system of states'. In 3ufendorf's wor! (e.g. =>A), this term, used interchangeably with 'confederate system' and 'association of states' (literally, a contractual 'society of states') applies to the wiss 6antons and the *nited 3rovinces (see Roshchin, 5/..), seen as an alliance or 'league' of states Lin which one 3rince or 2eneral of the Eeague e\$cells the rest of the 6onfederates, and is clothed with symbols of royaltyM (Hl.9). A loo! at 3ufendorf's (>./, p.A..) later formulation on the concept of a 'system of states' should clarify the point%

Then several tates are, by some special land (*sic*) so closely united, as that they seem to compose one lody, and yet retain each of them the overeiqn 6ommand in their respective Cominions; these we term *Systems of States* (O). +f *Systems* properly so call'd, these &wo kinds do especially fall under Botice. +ne, when two or more tates are sub9ect to one and the same <ing, the +ther, when two or more tates are lin!d together in one lody by virtue of some Eeague or Alliance.

A system, in 3ufendorf's view, is not the same as an 'irregular body'. Beither is it a unitary 'body'. ince the #mpire seemed to tend toward the second type of states' system, then it should function harmoniously as such for the sa!e of preserving its own e\$istence. -owever, the #mpire was not really a 'system of states'. It only had the potential of becoming one more easily than a unitary monarchy, if only states stopped to impose their wills on each other and balance against each other. It was still an 'irregular body'.

&he problem was that the #mpire as a body was Lattac!d by furious CiseasesM. +n the one hand, the #mperor wanted to centralise it. +n the other, the local princes demanded Lperfect EibertyM. As a conse"uence of such lac! of cohesion, their Lforeign alliesM had the unprecedented Lability to mould 2ermany to their own particular Interest and 7illsM (Hl.9). &he internal distinctions between each type of sub'imperial political unit also displeased 3ufendorf, Lso that 2ermany cannot even be regarded as a well'

ordered system of confederatesM (HII.@). As a prescription for reform, two things should be taken into account. First, Zufendorf did not believe in the existence of any concrete eternal threat to Germany, although the polity was certainly weakened in the multidirectional alliance-making between its members and foreign powers.^{=/} &thus, secondly, it would only be a matter of rearranging the #mpire into a system of states with the provision that the #mperor must not aspire to overaigntyM (*sic*) over the internal princes (HIII.G). As recent research on 17th century political rhetoric indicates, the Ottoman #mpire was the main candidate to an eternal threat to Christendom (Almond, 5/./, pp.1A; .=F; Wilson, 5//@, pp.>=ff). It would, therefore, make sense to undertake a more detailed analysis of its relative strength and potential courses of action. Zufendorf studies military and geopolitical capabilities of the 'threat' but concludes that 'there is no great reason for the Germans to fear them' (HII.G). Instead, the greater threat to the #mpire is the irregularity or 'monstrosity' of the political 'body' itself. As we shall see, Eeibni? begged to differ. -e did not see any problems in that 'irregularity' and arranged much of the argument around the notion that the #mperor was a leader of Christendom against what he perceived as the great Ottoman 'threat'.

The Empire as a union and state

For Eeibni?, the main immediate goal is 'the redefinition of the concept of sovereignty in a way which would allow the minor German princes to be treated as sovereignsM (Riley's commentary; Eeibni?, .9@@, p....). #mployed by the Cu'e of -anover as a kind of official historian, Eeibni? had the daunting task of justifying the Cu'e's right of legation independently of the #mperor's representatives.⁼ -anover's Lprimary problemM, writes a commentator, is that 'testphalia conferred sovereignty upon all those German rulers who had formerly been included in the -oly Roman #mpire; however, it had not abolished the traditional, essentially feudal, structure of the #mpire itselfM (<nutsen, .99>, p.9.). &he challenge for several middle-sized states in the #mpire was to obtain a right to diplomatic ties and representation while restraining the same access to even smaller political units. -aving a 'ticket to foreign policy' was a very important aspect of #arly (modern #uropean statecraft, not least because the institution

^{=/} Botice, however, that a few decades later Zufendorf would express concern with the hegemonic rise of France, especially after the #dict of Bantes (.=@A). A similar negative perception of France as a potential threat was voiced by other German Protestant authors of the period (Cauders, 5/./, pp.5.G'5F).

⁼ &he immediate goal was to secure representation in the peace negotiations in Birmegem (#ulau, .9G., p.=AA).

of diplomacy was strongly predicated on ceremonial precedence and its more substantive implications (Ilac!, 5/./, pp.GF'@G; Roosen, .9@/).⁵ In addition to this, another contextual element influencing Eeibni?'s writing was his political preference for a stronger role for the #mperor as Lthe secular arm of the universal 6hurchM (Eeibni?, .99@, p....). &he perception that 3apal and Imperial authority and leadership had been undermined by the 3eace of 7estphalia was a clear concern (<nutsen, .99>, p.9.). In fact, the philosopher too! up ecumenic reconciliation between 6atholics and 3rotestant as his personal pro9ect, with political implications (<enny, 5//=, III, p.>.). It was Eeibni?'s (.9@@, p...5) belief that Lone cannot refuse to 6aesar some authority in a great part of #urope, and a species of primacy analogous to the ecclesiastical primacyM. &his reflects Lhis constant emphasis on 6hristian unity in the face of the +ttoman threatM (Almond, 5/./, p.>). &herefore, the case for the rights of local members of the #mpire ought to be made in a careful manner, lest Imperial authority be challenged to an even greater e\$tent.

In order to accommodate both goals, the argument goes to great lengths before defining the -oly Roman #mpire as a sovereign state and, at the same time, a 'union' of sovereign 'regions' and other reasonably si?ed units with right to legation. <ey to this reasoning are, first, the differentiation between distinct political units and their 9uridical properties and, secondly, the introduction of a notion of sovereignty as a *divisi le* category. &he first type of political unit mentioned by Eeibni? (p...G) is the 'state', La fairly large gathering of men, begun in the hope of mutual defense against a large force, such as is usually feared, with the intention of living together, including the foundation of some administration of common affairsM. A 'city', in turn, Lwould seem to re"uire cohabitation such that the citi?ens can easily assemble when the call goes outM and, therefore, is much smaller as a political unit. A 'dominion' is Lan area of inhabited land served by a common administrationM and a 'region' is more or less similar, only much larger in si?e. A 'province', finally, e\$ists Lwhen a region is part of another still larger dominionM. i?e of territory, however, is not the only feature of differentiation between them. Although sovereignty also depends on territory, it is important to ta!e into account the 9uridical "ualification of this and other resources mobilised by a given polity.

@andeshoheit, or 'territorial hegemony', mentioned at the 3eace of 7estphalia (.=G@), entails, first, a '9urisdiction', that is, Lthe right of deciding cases or of handing

=5 &han!s to lain -ampsher'(on! for the 'tic!et' metaphor.

down judgments, and of coercing obstinate private persons (p...A). Secondly, it involves 'mild power of coercion', being able, when necessary, to use force on stubborn people. Finally, it means having the right to 'military might', when it is in the power of the one holding this right to assemble a military force which is sufficient for keeping the whole dominion in its duty. Territorial hegemony is the highest right of coercing subjects in this vein and encompasses full discretion to command all other things. A line is drawn, therefore, in terms of resources, between small polities with no 'diplomatic tie' and those middle-sized units with a right to legation. Territorial hegemony, being the highest right of forcing or coercing, is the kind of authority confirmed for units which exercise 'supremacy'. These are the larger powers which can wage war, sustain it, survive somehow by their own power, make treaties, take part with authority in the affairs of other peoples. Such right should be honored by the other major powers (O) as brothers and persons of equal condition (although, perhaps, of lesser power by a considerable degree) (pp...=>). To sum up thus far: despite the fact that there are several different polities in the empire, each with their own control over certain resources, we may still speak in a nuanced way about types of rights that each of these properties entail. Such is the method employed by Leibniz in order to argue that some of the middle-sized units, like Bavaria, were allowed to make their foreign policy while the smaller ones were still excluded from the privilege.

What about the relation between these units and the emperor? It is here that the creativity of the argument lies. Unlike what Althusius and Zueffendorf assume, says Leibniz, sovereignty is divisible, and is shared by both emperor and the polities of the empire. Several territories, he argues, can unite into one body, with the territorial hegemony of each preserved intact, provided each region keeps their right of having soldiers, making foreign treaties and the right to arms and conscription (p...>). Since the German princes can do all of these, their union into one body does not imply local loss of sovereignty. Each 'region' retains 'supremacy'. The only Roman empire is not a mere 'confederation' of sovereign states with no sovereignty of its own, but rather a 'union' and a 'state' itself.

A confederation is entered into by words alone and, if necessary, forces are joined. For a union, it is necessary that a certain administration be formed, with some power over the members; which power obtains as a matter of ordinary right, in matters of great moment, and those which concern the public welfare. Here I say exists a state (p...>).

In his assumption that sovereignty is divisible, Eeibni? manages to defend, at the same time, the ma9esty of the #mperor, influenced by the thin!er's desire to !eep 6hristendom under a single political head, and to ma!e the case for -anover's right to legation and foreign policy'ma!ing (,reire, 5/.., pp.5.'5). 1y implication, he also engages in sharp criticism of the traditional view of sovereignty as an indivisible property.^{=F} &here are several cases, li!e the wiss 6antons and the *nited 3rovinces, for e\$ample, in which empirical reality attests states in the form of 'unions'. Civeded sovereignty is, in fact, the rule and not an e\$ception, even in the most centralised states of that time, such as #ngland, pain and ,rance (Eeibni?, .9@@, pp...@'9).

&he notion of a union of states being more than 3ufendorf's 'system of states' or the traditionally accepted category of a 'league' or 'alliance' was, according to a commentator, a ma9or innovation in this contribution to the theories of federalism in the #mpire (Riley, .9>=, pp.5.; 5A).^{=G} Although there are many additional implications to this view of the #mpire as a union and a state (see #la?ar, .9@>, pp.@'./), it is clear that the move allows Eeibni? (.9@@, p....) to address some of his concerns with the 'bigger picture'. Arguing that the #mperor was not only sovereign in the full sense, but also Lthe born leader of 6hristians against the infidelsM, he also assumes that Imperial sovereignty allows the leader to constrain Lthose turbulent men who, without regard to what is permitted and what is not, are disposed to sacrifice the blood of the innocent to their particular ambitionM (p...5). Internal cohesion is particularly relevant for the sa!e of !eeping 6hristendom protected from the e\$ternal threat of the +ttoman #mpire. 6ontrary to 3ufendorf, Eeibni? saw in the +ttomans a very powerful and disciplined polity capable of doing great damage to 2ermany. &hey were not only 'infidels', as he writes, but also Lbarbarians, who err in whichever direction you might chooseM (p..5/). &he emphasis on sovereignty, right to control military resources and the statehood of the #mpire as a union was Eeibni?'s plea for the suspension of the internal differences as a condition for resisting e\$ternal attac!s. hortly after, the -oly Roman #mpire would indeed suffer from prolonged hostile relations with the +ttomans.

^{=F} L-obbes' fallacy lies in this, that he thin!s things which can entail inconvenience should not be borne at all 0 which is foreign to the nature of human affairs. I would not deny that, when the supreme power is divided, many dissensions can arise; even wars, if everyone holds stubbornly to their own opinion. 1ut e\$perience has shown that men usually hold to some middle road, so as not to commit everything to ha?ard by their obstinacy (O). In the 2erman assemblies, too, not everything is transacted by ma9ority vote, but some matters re"uire unanimity, all of which cases would seem anarchy to -obbesM (Eeibni?, .9@@, p...9).

^{=G} &he single relevant theorist to anticipate it was Eudolph -ugo, who had also wor!ed for -anover.

Metatheory in international political theory

I turn now to my analysis of metatheoretical elements involved in the views formulated by Althusius, Zuehlendorff and Leibniz on the early Roman Empire. My typology of metatheoretical research, set forth in part one of the thesis, refers to Intellectual and Contextual research on theory. These two types of inquiry combine, in turn, with two domains. They may be Internal or External to a chosen academic discipline. Most part of the analysis below focuses on the Intellectual side and is, not surprisingly, based on considerations External to the field of IR. Such limitation on possible uses of the types of metatheory is explained in at least three ways. First, I have deliberately selected a case pre-dating the academic discipline of IR in order to assess the impact of metatheoretical elements on theoretical investigation. The expectation for this 'control' case' is that metatheory indeed matters, although the lack of a well-defined disciplinary context does make a difference. For one thing, there is no way to analyse the *intra'* disciplinary elements of either Intellectual or Contextual analyses. The remaining possibilities, therefore, are those of metatheory focusing on Intellectual and Contextual points *outside* the discipline of IR. Secondly, there is only so much one can do in terms of Contextual research in the case of the material selected for this chapter due to the limited quality of the information on the immediate occasions of writing.^{=A} We are, thus, left with the Intellectual aspects which are External to IR as an institutionalised discipline. It is here that I concentrate the analysis. But, then, thirdly, this should suffice to illustrate the relevance of metatheoretical elements even in an ill-defined (or, rather, undefined) disciplinary context.

Before proceeding to the analysis of these three early (modern theories of the early Roman Empire in terms of their intellectual elements, let me explore as far as possible the Contextual side of the study. At least two 'drives' asserted some influence on the parameters of these theories of the Empire. First, there were immediate Contextual elements with an impact on theorisation. Althusius makes a clear apology of the Dutch resistance to Spanish rule as a relevant specific instance of just war and resistance against tyrannical rule. His practice as syndic in the Imperial city of Muenster, as well as his use of local examples further corroborate the link between the immediate context of writing and the way the Althusian theory of the legitimate limits of the 'public association' is set forth. In fact, Lhe intended to achieve for Muenster exactly what the

^{=A} Obviously, specialised commentators are expected to arrive at a considerably detailed description of at least the *inter* context of writing.

Cutch provinces had achieved. His political theory was meant to be the blueprint of political organisation (Fleeglin, 1999, pp. 105; G.). While Althusius was specifically interested in the dynamics of domestic resistance as a way of ensuring the harmony of 'symbiosis', Zuefendorf and Eeibni? wrote a generation later, and after the thirty years war. As a result, they were instead more concerned with the empire and its fragmentation. The tone of Zuefendorf's work is highly polemic, including the fact that it was initially published under the pseudonym 'Gronovianus'. It is said that this move enabled him, so far a relatively unknown professor whose fame stretched only to the local level, to navigate the waters between those defending the well-established absolutist centralism and those partisans of an accumulation of prerogatives in the hands of the estates (Fleeglin, 1999a, p. 105). This helps account for why Zuefendorf, disguised under a pseudonym, interrupts his argumentation on several occasions in order to contrast his view with alternative theories of the empire that he strongly rejects.¹⁰ As seen before, in the case of Eeibni?, he had clear directives from his employer that he should theorise the empire in such a way as to ascribe to it the right to legation in the peace of Westphalia (Fleeglin, 1999b). He had also a 'grand design' for a unified Christendom against the Ottomans. According to several legal scholars of that time, the dissolution of the empire (and not its permanence) would probably lead to peace and justice (Fleeglin, 1999c, pp. 105-106). Eeibni?, opposing them, seemed to be the only one who argued in favour of its continued existence and even emphasised its status as a political body and international actor (p. 106).

So much for the first contextual point. The second point is wider and has to do with the authors' religious backgrounds, which arguably influenced not only their specific writing on the empire but also their political theory as a whole. Religious background should be acknowledged as an extremely relevant feature to be considered in the analysis of ideas produced in the early Roman empire of that time, given the sharp religious distinctions triggered by the Reformation and the wars of religion.

¹⁰ The main one being that of Logislaw Ghemnit? (Fleeglin, 1999b), whose work on the empire was published right before the conclusion of the peace of Westphalia under the pseudonym 'Gronovianus'. Ghemnit? was employed by the Swedish crown, which had an interest in a weaker empire. He defended the supremacy of the lower magistrates and princes over the emperor (see Fleeglin's note in Zuefendorf, 1999, p. 105).

¹¹ France would be a key player in the conference and was refusing to recognise the right to legation of the smaller Imperial states. This also helps explain Eeibni?'s (Fleeglin, 1999c) parallel text on the same subject, written at the same time, but in French and in a pamphletary format as a fictitious dialogue. In this parallel text, he argues that the sovereign or potentate is he who is master of a territory powerful enough to become relevant in Europe, both in peace and in times of war, for treaties, arms and alliances (freely translated from Fleeglin, 1999c, p. 106). For a commentary, see Fleeglin (1999, pp. 106-107).

Althusius stands out as the only Reformed Christian (i.e., 'Calvinist') in the group. #mden was predominantly Reformed, but the Reformed faith was only officially recognised in the #mpire decades later, in the 3eace of 7estphalia. &he Reformed communities saw church discipline on doctrine (defined by their official confessional documents) and behaviour (also codified in the confessions) as *the* key to their identity (Wilson, 5//@, pp.F='@). LReformed Christians were kept in check by their local church consistories regarding both conduct and beliefs. If Althusius (who, later on, became himself a member of his local consistory) publicly practised or taught anything against the Reformed confession, he would be in troubleM (reire, 5/.., p.5G).^{=@} Calvin himself, as well as the Reformed confessions emphasise civil obedience as a normal state of affairs in Christian life and allow a small, but important, e\$ception for the case of 'tyranny'. &his certainly matters here, for Althusius (e.g. VVVHIII.>=) not only argues in detail about how damaging an e\$pansion of power by the ruler of the 'universal association' is to society, but also emphasises that resistance to tyranny must be 9ust, legitimate and organised under the leadership of 'lower magistrates'. 7riting decades later as Eutheran humanists, 3ufendorf and Eeibni? seem to be more comfortable in their e\$ploration of controversial ideas. &heir theories have a more secular tone. -owever, both indicate the desire to see stronger political coordination in Christendom. -aving worked in -eidelberg, another predominantly Reformed city, 3ufendorf (5//5) became more aware of details in Calvinist theology and tried to reconcile several points of tension between the Reformed and the Eutheran confessions. -e argued for a unified 3rotestant church in the #mpire comprising both Eutherans and Calvinists.⁼⁹ Eeibni? provided Lsharp criticismM (Qurbuchen, 5//5, p.\$vi) and adopted an even more 'ecumenical' stance, defending a unified Christian church and polity (see CRren, .995, pp.5/A'./). -e was a Eutheran Lwith Catholic sympathiesM, re9ecting Lanti-CatholicM political theologies that were so popular in 3rotestant circles (Almond, 5/./, p...).

uch 'drive' behind his theory of the #mpire is ascribed to his Lpersonal theological' political ideal of a reunified Republic of ChristendomM more than anything else (Bi9man, 5//G, p.GF). &hus, religious bac!ground is another relevant 6onte\$tual factor

^{=@} Althusius was successful keeping his theory within confessional limits. (urdoc! (5//G, p.>F) confirms the strictness of church discipline in #mden, and 7itte (5//>, pp..A5; .99) registers .=> as the date for the election of Althusius as a church elder, after several editions of *Politica* were widely available and their contents verified. (oreover, he was not re"uired to stop writing, having published *<icaelogicae* (.=.@), his theory of 9ustice, during his position as an elder.

⁼⁹ Qurbuchen (.99@, pp.G.9'5.), however, believes 3ufendorf's 3rotestant inclinations and anti-Catholicism should not be overstated or become the controlling factor in the interpretation of his theory of the #mpire.

shaping these theories.

Having dealt with the *Contextual* elements that may help us understand some of the features shaping theoretical argument, I turn now to a more detailed analysis of several ways in which these theories are *Intellectually* shaped by metatheoretical principles. These discursive mechanisms are either implicit or explicit. On the 'implicit' side I detect two types of metatheoretical 'impulse': the issue of 'parts and wholes' and the issue of 'how to theorise'. On the 'explicit' category I highlight the intellectual influences and authorities acknowledged by the authors themselves (and to whom they react in their work(s)), as well as efforts of 'bridge-building' between their broader philosophical positions and the specific theories of the early Roman Empire.

On the first implicit element – the issue of parts and wholes – it is clear that each of the authors relies upon implicit assumptions on the nature of reality applied to their object of analysis. Zuzendorf and Leibniz adopt a more individualistic approach than Althusius. They understand the Empire based on their view of the interacting parts. In fact, for Zuzendorf, a key problem with the Empire as a political 'body' is that its parts often interact in a conflicting way. This prevents it from becoming a 'system of states'. Although the hope for a strongly centralised Empire is gone, Zuzendorf still expects several of the problems to be corrected if only the contracting states begin to interact like a proper system. A system of states is not simply a set of interacting parts. The set itself may be either a 'monstrosity' or a 'system'. It is, rather, the quality of the interaction of the parts that affects what the Empire becomes as a whole. Leibniz, in turn, not only believes that some irregularity in the internal relations of a polity is empirically the norm (rather than exception), but also argues that the Empire is in itself a state, and more than just the sum of its parts. In his scheme, Germany is a union of territories, and not a mere confederation or 'system'. His idea that a union of states with a permanent administration constitutes more than a mere 'system of states' is a key element to later notions of federal states. A more unique view is adopted by Althusius. He spends a great deal of time distinguishing between the inner qualities of each type of association. While we may speak of villages, cities and provinces in terms of a certain hierarchy of parts and wholes, we must not read a 'universal association' like the Empire as a set comprising every single association in its territory. Strictly speaking, professional guilds, churches and other associations which do not have a 'vocation' to promote public justice with the 'power of the sword' are wholes *in themselves*, distinguished according to their respective 'vocations'. For example, the (religiously

"ualified) local congregation is a part of the 'glassis' and then 'ynod' of the 'hurch', an association with similar "ualification. It is not a part of the '#mpire' as a polity (E. 2. ,reire, 5/.). In a mathematical figure of speech for each conception of the #mpire, if 3ufendorf's 'keyword is 'addition' of parts, the Eeibni?ian operation consists in 'integration' into a larger whole and the Althusian counterpart would be that of 'differentiation' of several wholes.

Bot only are there concealed assumptions about the ontological relation between parts and wholes, there are also implicit principles on how to theorise. ,irst, there seems to be a 'hidden' presupposition about what one should make of 4ean lodin's (.AF/' .A9=) notion of sovereignty as an indivisible property, a highly prestigious argument at that time. Althusius and 3ufendorf show 'due respect' to lodin's theory when they cite it, although their conclusions are considerably different to those of lodin. Eeibni?, in turn, is rather bold in his introduction of the notion of a relative view of sovereignty, re9ecting the common use of lodin's concept in #arly (odern political theory and enabling new possible claims about federalism and divided sovereignty.

econdly, in their 'way of theorising' the authors implicitly reflect metatheoretical directives provided by two 'schools of thought' of their time. Althusius was a follower of the rules of scholarship established by the ,rench Reformed logician 3etrus Ramus (.A.A'.A>5). &he 'Ramist school' had a widespread reputation (3lett, 5//G, pp.A=>) and its influence may be distinguished in Althusius by the schematic organisation of the arguments in categories, from the most general to the most particular.[>] In fact, his boo! contains diagrams 0 another Ramist trait 0 indicating how each sub9ect addressed in the theory is a branch of a more general tree, until we arrive at the most basic category, that of 'politics' itself (remember ,igures A.. and A.5). &hus, a common criti"ue of his style has to do with this Lrather tiresome methodological dichotomi?ing and hairsplittingM (-ueglin, .999, p..9) which was, nevertheless, rather usual among other Reformed scholars.[>] &he other two writers were under the influence

>/ I "uote from the first #nglish edition of the @ogi%e, dating from .A>G (VH, pp.9G'A). L&he chiefe e\$amples of the methode are found in artes and sciences% in the which although the rules be all generall, yet they are distinct by there degrees (*sic*)% for every thing as it is more generall is first placed. &he most generall therefore shalbe first placed% the ne\$t shall followe these which be immediatly c'tained under the general, every one orderly unto the most speciall which shalbe last disposed. &he definition therefore as most generall, shalbe first placed% ne\$t followeth the distribution, which yf it be manifold, and of divers sortes shalbe first divided into his integrall partes, ne\$t into his formes and !indesM. ,or Ramus (pp..//'), brea!ing these rules by Lhaving some degrees of the order invertedM would LmutilateM the method and Ldeceave the auditorM, which is Lpreposterous and out of all good fashion and orderM.

>. Another well!nown e\$ample can be found in the writings of 3uritan theologian 7illiam 3er!ins (.AA@'.=5).

of a mathematical version of humanism, and this entails a series of implications for their 'way of theorising', not least following the *more geometrico* the principle that an ethical system should be deduced from basic axioms imitating Euclidean geometry.⁵ The popularisation of this particular kind of secular and geometric humanism has been portrayed as a reaction to the political, religious and intellectual problems emerging from the 17th century crisis that culminated at the Thirty Years War (Koulmin, 1997, pp.175). Zuefendorf's test is described as having a strong philosophical subtlety in the mutuality of theory and practice, a strong empiricism or realism, and opposition to scholastic categorisation and argument (Eidler, 1997, p.1). He employed a hypothetical resolutive compositive method (Louchet, 1997, p.14), in which propositions about the empire are derived 'from below', from the nature of states, which, in turn, are derived from his argument on human nature (Mullau, 1997, pp.14-15). As a commentator puts it, his thought follows a proposal to regulate relations among men, things, persons, communities and their conditions and particularities according to mathematical principles (Wesbe Elanos, 1997b, p.14). He attempts to derive the conceptual construction like an edifice of propositions derived from basic axioms (Wesbe Elanos, 1997a, p.14). A similar metatheoretical 'impulse' can be seen in Leibniz. His argument for the legitimacy of a strong and sovereign empire is predicated, via a long chain of reasoning, on his notion of justice, later systematised in terms of his 'universal jurisprudence' (Birman, 1997, p.14). The 'way of theorising' justice in his political thought posed a dilemma clearly influenced by the *more geometrico*. On the one hand%

Throughout his life Leibniz was tempted to assert that principles of justice, as 'eternal verities', had the same status as A or A⁵G, and for an obvious reason% one of his great hopes was that of reducing all complex propositions to their simplest form, to primary and irreducible concepts whose predicates were clearly contained in their subjects (Riley, 1997, p.14).

However, on the other hand Leibniz was also worried that justice, which had to be promoted by true statecraft, if defined strictly in this mechanistic way, would allow little room for human agency. Therefore, despite trying to define it simply in terms of harmony, or proportion, of ratios as precise as any in mathematics he also came to the conclusion that a dynamic aspect of action should be added to it (pp.14). If we look

⁵ Wesbe Elanos (1997, p.14) traces Zuefendorf's choice back to his interaction with the ideas of mathematician Richard Weigel (1541-199), who transposed the mathematical demonstrative method to (oral) philosophy.

at these implicit 'metatheoretical' principles such as the writers' dialogue with Iodini's theory and the 'schools of thought' that they adopted, we are further able to understand the nature of the diversity of these theories speaking of the same object.

The implicit metatheoretical elements driving these three views of the -oly Roman #mpire in #arly (modern political theory that I have selected for analysis have to do with underlying assumptions about parts and wholes and about how one should proceed in theoretical thought. Turning now to what is clearer in the texts themselves, a first point relates to the Intellectual influences shaping each respective argument. There is, for example, the role played by the biblical text. Citations abound in the work of Althusius, who not only employs them as authoritative sources and normative perfection (taken in context), but also as referring to historical illustrations of his own points.^{>F} Zuefendorf (HI.=) demonstrates great knowledge of the bible, but opts for not deriving much argumentation from that source. Instead, we can notice, for example, his negative use against contemporary monarchists who defended the idea, based on their interpretation of the 100! of Daniel (>..'5@), that the #mpire was strong enough to last until the end times.^{>G} Biblical citations also have a low profile in Eeibni?. However, unlike Althusius, and with Zuefendorf, he highlights his spite for contemporary colleagues. He says that, among his peers, he lacks the aid of good writersM (.9@@@, p...G). Zuefendorf, as well, denounces his peers for 'the careless compilation of others' opinions as a 'new boom!'M (HI..) and for rushing to comment on public law with little or no knowledge of civil affairsM (HI.F). Each of empirical or practical knowledge is also an objection raised by Eeibni? against them, who have only eyes for what is ancientM and whose experience Lsupposing they have any 0 does not go beyond the gates of the tribunalsM (.9@@@, p...F). This is important, because both him and Zuefendorf avoid building an argument based on pre-established conclusions of deductive scholasticism, except where they find them relevant. Thus, for example, Eeibni? (p...A) cites a dictum by Ialdus de *baldis (.F5>'.G//) on sovereignty.^{>A} Some are pulled by

>F In fact, the subtitle of his work is L3olitics... illustrated with sacred and profane examplesM. 'sacred' meaning 'taken from biblical narratives'. Althusius evaluates 'the polity of the JewsM of the +ld &estament as being as near perfection as it could possibly be, although he was controversial for defending the idea that different circumstances of his day would require distinct applications of the same +ld &estament principles (#la?ar, .99=, p.F.G).

>G Daniel's prophecy alludes to a kingdom that will last forever. Issue is taken against Cietrich Rein!ring! (.A9/'.==G) in particular, who had defended the idea that the -oly Roman #mpire was the eternal continuation of the Roman #mpire. But notice Zuefendorf's mode of argumentation, laying more emphasis on historical information than biblical. He cites -ermann Bonring's (.=/='.=@.) thesis that the introduction of Roman law into the 2erman #mpire was fairly recent.

>A According to Eeibni? (.9@@@, p...A), Ialdus used to say that hegemony inhered in a territory as the mist to a swampM.

Althusius' much friendlier reception of contemporary works and classic Roman Catholic scholasticism (C6oy D la!er, .99., pp.AF'A; 7itte 4r., 5//@, p..AF). -owever, this is not necessarily to say that he himself subscribes to the scholastic method. Instead, his adhesion to the Ramist intellectual movement better accounts for his critical engagement with contemporary and past authors of a different persuasion. Following Ramus' (.A>G) advice, he cites them where they agree with him, in order to make the point that the thesis can be maintained even by those who start from distinct assumptions.^{>=} In these traits, we can see yet another aspect of how the authors follow their respective 'schools of thought'.

Besides drawing on a clear dialogue with, and antithetical counterpoint to, several intellectual authorities, the theories of Althusius, Zuefendorf and Leibniz were also formulated by bridging between their respective philosophical positions as a whole and specific political theory in this particular case. Again, the Ramist principles followed by Althusius were extremely relevant starting points. His discussion of the Empire is only an illustration of a general treatise on politics, where he tried to defend the subject as an independent discipline. A key rule of scholarship established by Ramus was that each art or science has its own purpose and that anything beyond a given field should be parsimoniously excluded from the investigation in that field (Garney, .99A, p.vii). Althusius fully endorsed this principle. In fact, it became the backbone of his contention that political science had to be established as a separate discipline (Wiegman, .999, p.>G). Another rule requires one to place an argument immediately next to the category to which it belongs. Following this metatheoretical directive, Althusius finds in 'symbiosis' the most general starting point of political science. The unique methodological contribution here (O) lies in an unprecedented determination of the political (Wiegman, .999, p.>9). The author was quite aware of this, having reviewed and dismissed alternative starting points available at the time.>>

Although this 'bridge' between philosophy and special theory in Althusius requires extra work on his intellectual background, in Zuefendorf and Leibniz the 'bridge'

>= In his theory of 'logic', Ramus (.A>G, VVII, p.===@) has a prominent place for citations of *Testimonie humaine*, including famous sayings (*sic*) and sayings also of wise men. Says Ramus% Lo Christs himself, the Apostles, and Evangelists do confirme their doctrine, by the lawes of Moyses% the Physicians, by the auctoritie of Hippocrates, Dioscorides, Galen, and such others% the lawyers, by Iustinian% and the mathematicians, by Euclides. & this is a clear Ramist trait, and an example of the novelty of Ramus' approach to logic compared to his contemporaries (<enny, 5//=, III, p..5).

>> Part of this decision also involved including private associations (families, for example) under 'political science', when they were sharply separated from the subject (Althusius, III.G5).

is easier to identify. 3ufendorf's 'monstrous' #mpire is an instance of 'irregular polities', which in turn are degenerations of the 'civic body'. &his thin!er is well !nown for his definition of the personality of states as not merely a legal fiction, but rather La real autonomous moral person with the capacity to will, deliberate and pursue purposesM (loucher, 5//., p.A=>). overeignty is, in 3ufendorf, Lthe soul that animates the person of the stateM (p.A=9). &he state, in turn, emerges in conformity to the law of nature% human beings promptly realise their natural impulse towards cooperation. &hey also re"uire safety and enforcement, which is Lbest achieved when bac!ed by the power of the stateM (p.A=A). &herefore, an 'irregular polity' with a 'sic!' soul ultimately prevents human beings of unfolding their full potential in cooperation. 3ufendorf's criti"ue of the 'monster', in short, is closely lin!ed to his broader philosophical views (see also -uesbe Elanos, 5//9, p.GFF).

,inally, we can notice a similar 'bridge' in Eeibni? as well. In his hands, the notion of sovereignty, for e\$ample, suffered Lan e\$treme downgradingM due to his philosophical Lemphasis on charity, welfare, and reasonabilityM elsewhere (Riley, .9@@, p.5=). 7hile he clearly ac!nowledges the relation between power and sovereignty, he also noted that power must go together with reason. &his con9unction 0 as opposed to its separate parts 0 Lis not only the foundation of beauty and 9ustice, but of true statesmanshipM (p.5G). Indeed, LnothingM in his political theory Lis more important than benevolenceM, which he connects to '9ustice'. ,or this reason, he urges wisdom and virtue from rulers, who should Ldevote all their efforts to the public welfareM (p.5A). (oreover, his view of a state (and, by implication, a union of states li!e the '#mpire') was e\$tre mely individualistic. &his is also a clear conse"uence of his philosophical assumptions, given that his metaphysics only allows for the reality of the individual (p.5=).>@ As a result, unli!e 3ufendorf, Eeibni? did not ascribe 'legal personality' to the state (2riard, 5//>). &o sum up% another metatheoretical element that 'drives' and differentiates their theorising is the fact that, in these studies on the -oly Roman #mpire, Althusius, 3ufendorf and Eeibni? bridge between their respective philosophical formulations and their specific use in political theory.

)inal remar*s

&he main concern in this chapter was to provide the first illustration of a series on how metatheoretical elements may shape the study of world politics. ,or this initial analysis,

>@ ee <enny (5//=, III, p.>5) for a commentary.

three early (modern theories on the early Roman Empire have been selected, expounded and then compared. In harmony with the typology of metatheoretical research introduced in the previous chapters of this thesis, I have based the analysis on contextual and Intellectual 'drives' behind the formulation of each of those theories. My analysis of the contextual elements, accidentally constrained in depth by the "quality of information on the immediate contexts, still yielded a number of points about the occasions of writing and the religious background of the writers. On the Intellectual side, I have noted implicit and explicit metatheoretical elements. The first implicit 'drive' identified was that of how each of the authors relied on ontological assumptions about wholes and parts and applied them in their theories. The second 'drive' had to do with the 'ways of theorising' their reaction to an implicit 'rule' of that time that Iodius's well-established concept of sovereignty should be followed, and the 'schools of thought' that shaped their theories. Generically speaking, these are views on 'how to do political theory'. Both implicit intellectual 'drives' help us account for why these theories speak in similar terms about the same object and yet are so different. This, however, was also complemented in this study by an analysis of two explicit intellectual 'drives' behind each theory. The first 'drive' clearly present in the texts is the "quality of the interaction with classical and scholastic authors, including biblical material. Their respective views of intellectual authorities, in turn, may also be traced back to each author's subscription to a certain school of thought. The second 'drive' relates to the way in which the writers bridge between their general philosophical theories and their specific theories of the early Roman Empire.

The purpose here was to illustrate the relevance of metatheory with reference to a historical conceptual debate that predated the establishment of an institutionalised academic discipline of IR. This initial example was specifically designed to address the concern that everything I have been claiming throughout this thesis is contingent upon the existence of a well-established academic discipline. In response to this possible objection, I have highlighted several ways in which contextual and intellectual 'drives' have shaped pre-disciplinary theories of world politics. The differentiation between theories focusing on the same object of study occurs, to a large extent, in function of these contextual and intellectual 'drives'. I must, however, clarify the nature of my response. First, I do not claim that the impact of such 'drives' has a similar nature as it would have in a disciplinary context. The fact, for example, that it is not possible to apply the distinction between metatheoretical analysis inside and outside the field,

stands out as a key missing element in a pre-disciplinary or non-disciplinary case. In order to further clarify this qualitative difference, I suggest in the next chapter that a comparison between contemporary IR theories of the transformation of the early Roman Empire and the European states' system could yield relevant results also with reference to the distinction between Intellectual and Contextual elements Internal and External to the field. The first qualification, then, is that although metatheory may still matter in the absence of a clear disciplinary context, we should not expect its impact to be of the same quality as that which occurs in an institutionalised discipline. Secondly, by Intellectual and Contextual 'drives' that shape theories in different ways, I do not mean a deterministic impulse, but rather a 'constraining and enabling' effect. The fact that Althusius was a Ramist shapes his discourse in a certain way, but another Ramist could have theorised the same Empire following the same metatheory and arriving at a different conclusion. This is sufficiently clear in this study when we notice differences between Zuefendorf and Leibniz, both following the '*more geometrico*', or same generic 'way of theorising'. Thus, the argument that metatheory *shapes* theory should not be confused with the stronger claim that metatheory *determines* theory.

Apart from completing the first illustration of roles played by metatheoretical elements with this case, a 'spillover' result emerges, which should particularly interest research communities in IR and historical sociology. In pursuing the goal of highlighting roles played by metatheory in the historical conceptual debate selected for analysis in this chapter, my own procedure has made use of metatheory at another level. IR scholars *metatheorise* in this way, when they study similar historical conceptual debates in light of the Contextual and Intellectual elements shaping the theories that they analyse. In my own scheme presented in part one of the thesis, this would be a *hermeneutical* and/or *historical* 'way of metatheorising'. Political sociologists, in turn, might be interested in how early modern ideas on world politics shaped the practice of world politics and could, therefore, benefit from a *metatheoretical* analysis of these ideas. In this case, however, the Contextual study would have to be reversed to cover the impact of theories on their setting, instead of what I have done in this study (i.e., context to theory). In my typology, this would be metatheory as a 'tool' of social theory, and could be employed in a *critical* way. There are two key conclusions here. One is that metatheory may indeed have an impact on theory even in a weak or non-existent disciplinary context. This impact is distinct from the influence it would have in an institutionalised discipline. We shall see several differences between one configuration

and the other in the next couple of chapters. The second conclusion is that this study of historical theories, being itself a metatheoretical study relevant to IR scholars and political sociologists, illustrates at another level how metatheory can help their field in IR by shedding light on theories and their contents.

Chapter 2

Metatheory and Theoretically Oriented History in IR1

Narratives on the Peace of Westphalia

Introduction

The Peace of Westphalia (1648) is one of the most cited historical events in IR, often with reference to the role it allegedly played in shaping the current international system. The discipline subscribes to a traditional narrative that frames Westphalia as the 'big bang' of modern world politics. The 'classic' account, still found in textbooks, can be traced back to its initial stage of formation in close dialogue with international law.⁹ It basically involves the idea that the treaties of Münster and Osnabrück (1648) at the same time fragmented the only Roman Empire into a set of independent and autonomous states and undermined papal authority in Germany. The new standards of world politics laid down by the settlement allegedly prescribed a broader society based on the principles of non-intervention, secular international law and an anti-hegemonic balance of power. From Europe, these norms would, then, be imposed by the great powers on their colonies or otherwise absorbed by the periphery of the international system, thus leading to a global society based on universally shared norms of sovereign statehood of the world we have today (see Çayaoglu, 2005, pp. 97-98). (More recently, however, a revisionist body of literature has emerged in IR which challenges the assumption that this single event in history accounts for the constitution of modern world politics. Although it lacks theoretical uniformity, the new 'revisionism' shares on its diverse fronts the impulse to critique traditional historical claims on the role and relevance of Westphalia. Studies range from factual clarification (Barvalho et al., 2005; Gasner, 1999) and outright 'iconoclasm' (Leaulac, 2006; Siander, 1999, 2005) to a 'qualification of the Westphalia narrative embedded in wider theoretical claims

⁹ For classical statements, see for example, the oft-cited paper by Cross (1964), Morgenthau's (1964, pp. 1-5; 5-7) textbook and -er? (1969) widely read study. For more recent works, see for example Watson's (1995, p. 1) treatise on diplomacy and -eld's (1994, p. 5) use of the phrase 'Westphalian model'. (Most of this literature is self-referential and advances claims sustained almost exclusively by secondary sources. For a recent overview of the classical narrative and its transmission, as well as new material found in recent textbooks, see Carvalho *et al.* (2005, pp. 97-100) and Schmidt (2005).

(-avercroft, 5/.5; trauman, 5//@).

In this chapter I analyse three well-known texts on Westphalia. Adam Watson's *Evolution of International Society* presupposes the classic story and incorporates it into an English school comparative study of the operation and modification of several international orders in history. Aware of how IR has challenged this traditional account, Daniel Philpott attempts to retain some role for Westphalia in his *Revolutions in Sovereignty*. Philpott's 'soft constructivism' portrays the event as the closest we can get to an all-encompassing 'revolution' leading to the implementation of many of our contemporary principles in political practice. In his view the 'conventional wisdom' on Westphalia can still be maintained. Immanuel Kant, on the other hand, draws on a contrasting theory that asserts the primacy of material relations. He sharply criticises what he calls *The Myth of 1789*, defending instead the notion that, if it led to any changes at all, Westphalia would be a marker for the consolidation of a pre-modern dynastic states' system, not a modern international society. In my analysis of these three key IR studies on Westphalia, I advance two claims. The first is that such distinct accounts of the same event relying upon the use of different theoretical approaches are shaped by distinct metatheoretical 'drives'. For example, the assumptions indicate why the relation between theory and history in IR cannot be addressed in pure 'factual' terms alone. The second claim is that, besides shaping historical/empirical research on the role of Westphalia, metatheory also plays a role in our study of these narratives (including my own), and can help address a number of issues pertaining to the theory/history complex in IR scholarship. For most part of this chapter, though, I look in depth at each of the selected studies and their metatheoretical assumptions. Following that, I address the second concern more briefly, and in light of some of the IR literature critical of the traditional story on Westphalia.

'Westphalia': legitimate authority and anti-hegemony

Watson's approach draws on earlier English school theorising in Martin Wight's (1992) *Systems of States* and in Hedley Bull's (1977) influential book, *The Anarchical Society*. It also incorporates some of the ideas introduced in a volume edited by Bull and Watson (1986) on *The Expansion of International Society*. The aim is to compare several states' systems in history, asking whether (and how) they have influenced our contemporary system (Watson, 1995, p.). In other words, a system operates when political units are sufficiently involved with one another (p. 10). According to Watson (1997, p. 10),

order and authority in states' systems can be understood in terms of La notion of a spectrum between absolute independence and absolute empire, where both extremes are ideal types, or theoretical absolutes, and do not occur in practice. Fully independent states are the ultimate authority defining domestic and external policy. #ven though, ideally, they would be guided by individual calculation, practical pressure and constraints of systemic interaction eventually lead them to consider the system as a whole (Watson, 1995, pp. 1-5; GA). They may decide that hegemonic authority is beneficial in the sense of providing a framework for a smoother operation of the system. Hegemony exists when a political unit (or a group of powers) leaves others internally independent but has some control over inter-unit relations and the operation of the system (1995, pp. 6-7A). Dominion, in turn, involves imperial control over both external and domestic affairs of other units to an extent, although they also retain some of their autonomy. The last category of empire is direct administration of different communities from an imperial centre (1995, p. 8A). All international systems are organised in terms of these patterns of hierarchy.

*nderstanding how such arrangements operate and change in each system is important because it helps us compare and contrast historical cases. *nlike Watson's earlier framework (see Watson, 1995, pp. 9-10), which assumed a clear-cut distinction between regular and suzerain systems, Watson's (1995, p. 1A) spectrum allows for a wider range of relations of authority and subordination in world politics. u?erainty, La vaguer concept, is only one among these other markers on the spectrum, and involves overlordship that can be formally agreed upon by others in the system or merely by tacit acquiescence. (oreover, he also stipulates that international systems tend to a hegemonic organising principle on the spectrum (1995, pp. 1A-2). -e portrays the spectrum as an arc, with its midpoint at the bottom of the pendulum's swing, somewhere between hegemony and dominion. &his is, of course, a metaphor for describing the tension between the desire for order, more easily attained via hegemonic authority, and the desire for independence, which contests excessive accumulation of prerogatives by hegemonic powers. If we pay close attention to the way the pendulum swings in each historical case, we may obtain a more dynamic conception of the operation of international systems. But patterns of order and change in hierarchy are just one side of the theory. The other side is reflected in Watson's concern for the role of community and culture in international order. Shared values, rules, custom and other elements are normally integrating forces in a system. ly being part of a

community, political units will tend to operate within the same core institutions and parameters, going beyond arrangements of convenience in merely procedural matters (.995, pp.='.>). (oreover, there are interplays between community and ordering principles. ,or e\$ample, when a powerful e\$ternal actor enters into a sustained pattern of hegemonic interaction in the system, the cultural clash between the hegemon and the others considerably affects the system as a whole. Another e\$ample is the role played by shared cultural norms and practices constituting what counts as a legitimate relation of authority in the system.@/

7atson's theoretical framewor! becomes clearer as he proceeds with the empirical comparative analysis of states'systems. ,ollowing a series of studies of world politics in the ancient world, he raises Lprovisional theoretical deductionsM. ,or e\$ample, systemic involvement in La web of economic and strategic interests and pressuresM leads to the development of Lsome set of rules and conventionsM regardless of a shared cultural bac!ground (.995, p..5/). If, on the other hand, there is such bac!ground, then we should e\$pect it to deeply affect the institutional outloo! of the system, as well as its movement across the spectrum of authority, thus pushing shared rules beyond a merely procedural status, in order to include Lshared values and aspirationsM (pp..5.'5)! ystemic stability, especially when an hegemon manages the system, involves more than a Lbalance of material advantageM. It also re"uires some level of agreement on the legitimacy of the pattern of organisation (pp..F/'5, see note =).@ 3reparing the way for his discussion of 7estphalia, 7atson (pp..F@'A/) establishes #urope's cultural coherence by loo!ing at norms and institutions in medieval 6hristendom. -owever, these Lwere not devised to manage the pressures of a systemM, for rule and interaction was Ltoo diffusedM and Ltoo localM, thus preventing us from defining that community as an international system (see 7atson, .995, p..A. note =). It was rather the impact of renaissance humanism on political culture and practice 0 and its diffusion from Italy to the rest of #urope 0 that turned it into a system of states (7atson, .995, pp..A5'=@). Its position on the spectrum was close to 'multiple independences', and there were evidences of an emerging balance'of power rationale.

-owever, the *de facto* power e\$erted by #uropean absolutist rulers still re"uires

@/ L&he rules and institutions and the accepted practices of a society of substantially independent states need legitimate authority to ensure habitual complianceM (7atson, .995, p..>).

@. ummarising some of his preliminary conclusions, 7atson postulates% LEegitimacy pulled the management of a society of states or communities towards the point along the spectrum where the communities concerned felt most comfortable. In so far as it influenced the way a society was managed, it was usually a force for stability and continuityM (.995, p..F/).

stability and legitimacy. To complicate things further, argues Watson (1995, pp. 9-10), the Protestant Reformation divided Europe into two religious factions, each with their respective political project. Loyal to Roman Catholicism, the Habsburg Emperor attempted hegemonic rise. Protestant polities, which were not originally directed against hegemony in the state system, became anti-hegemonial because of the commitment of the Habsburgs to the Catholic side. In Germany, many local princes combined feudal customary law with Lutheran aspirations to resistance. The 16th century witnessed violent wars of religion in the Holy Roman Empire, until the Peace of Augsburg (1555) placed the power to determine official religion in the hands of local princes. Catholic France, however, joined the anti-hegemonic side against the Emperor, thus stressing the desire for a balance of power pulling the system towards the 'multiple hegemonies' end of the spectrum. This pattern, however, still reflected no 'grand systemic design' and, therefore, no consciously anti-hegemonial international society (p. 10-11). New wars of religiously based alliances emerged, in particular the Thirty Years War. Indeed, it was the Peace of Westphalia that put this pattern to an end and provided systemic legitimacy to anti-hegemony as new organising principle, turning it into the constituent legitimacy of the European society of states. Moreover, the settlement legitimated and standardised the Renaissance practice of independent and sovereign statecraft, previously adopted on an *ad hoc* basis (pp. 11-12). The systemic pattern was reconfigured from rigid leagues and alliances to anti-hegemonic legitimacy. For the next few centuries, common diplomatic practice in Europe would involve, like Westphalia, 'congresses' of representatives of several polities after major wars and frequent shifts in the formation of alliances.¹⁵

The Evolution of International Society heavily relies upon ideas advanced by the British Committee on the Theory of International Relations, the early English school's main forum of discussion. Watson himself was a key figure in the Committee and worked closely with Martin Wight, Herbert Butterfield and Hedley Bull in the development of a unique view of the link between history and international theory (Cunne, 1990). One of the metatheoretical principles of the Committee was that theory is a product of practice and that the history of practice should be employed to develop theory. The group drew on the assumption that in order to understand the contemporary world we need some sense of how international societies developed and operated in the

¹⁵ On post-Westphalian congresses of great powers following an English school perspective, see Blair (1999).

past (Luttwak and Little, 1999, pp. 15; 16). The main source of these principles for the Committee, in turn, was the work developed by 19th century scholar A. E. H. Green (1845-1915). Watson develops his formulation in dialogue with the research output of the Committee under the inspiration of Green's historiography and conceptual framework.^{@F} In other sources, Watson is eclectic and avoids too much theoretical discussion, IR texts included. Although Watson adds his own contribution to the framework of the English school, most of the metatheoretical drive of the work can be traced back to the principles advanced by his colleagues in the Committee. We may speak of at least two sets of directives in this regard. First, there are points referring to the nature of historical inquiry and its relation with international theory. Secondly, there are elements of a view of the nature of social interaction and its historical formation in terms of order and change at the international level.

In the first set of metatheoretical directives of Watson's view of the relation between theory and history we can identify at least three points. To begin with, there is the issue of historical parallels and comparison. Providing a comparative and historical analysis of states/systems in the past with a view to how they helped shape our current world is, in fact, the key aim of *The Evolution of International Society*. This concern is specifically visible in the discussion on Westphalia. Notice, for example, how Watson (1995, p. 9) concludes his evaluation of the role played by smaller polities emerging from the fragmentation of the early Roman Empire:

(more than a hundred small principalities (O) were invited to the negotiating table and acquired a sacrosanct quality of sovereignty (O). They all participated independently in the diplomatic dialogue (O). The concept of independence for a similar multitude of small states in our present international society, formed from the fragmentation of empires, and their presence at the permanent congress of the United Nations, has evolved from the Westphalian settlement and bears an inherited resemblance to it.

The teleological and comparative character of this passage is clear in the historical parallel drawn between the Westphalian order and the post-colonial order established in the 19th century. This illustrates the first point pertinent to the relation between theory and history in the work.

Secondly, by building analysis based on historical parallels, Watson (1995, p. 9) indicates the inductive and empirical character of his intellectual enterprise,

^{@F} In Watson's work, Green's name is mentioned several times not only in connection to 19th century great power politics, but also to the theoretical notion of a 'states/system' (Luttwak and Little were impressed by Green (...)) (Watson, 1995, p. 9).

although research will always require a theoretical approach. Despite the comparative inclination, the author is careful enough to require, first, a contextualised understanding.

States' systems need to be analysed in their own individuality and on their own merits before we compare them. Then we look at Europe before Westphalia, for example, we must look at its unity and individual nature, for only in that way can a historian understand it and perceive its original features (p. 14). However, we should bear in mind that Watson is not and, before him, so were Lutterfield and Wright are unimpressed with merely descriptive empirical study (Luan D Eittle, 1999, pp. 11-13). For one thing, information needs to be organised and employed within a framework. And thus, we are advised to limit our attention to those aspects which concern the relationship between different communities, and which have contributed to our present principles and practices when we research the complex array of factors in the European trajectory between the Middle Ages and Westphalia (Watson, 1995, p. 14). For another, beyond the particularities of each historical state's system, we need general categories that will allow us to compare them. Watson's constant reference to legitimacy and the spectrum between empire and multiple independences is a clear indication of this. Most of the systemic change ascribed to Westphalia in *The Evolution of International Society* is that the settlement crystallised a type of dynamic alternation between legitimate multiple independences and illegitimate attempts at hegemony on the spectrum.

In connection to this, the third point on the relation between theory and history has to do with how this relation occurs in concrete terms with reference to the early English school. Watson derives most of his framework of analysis from Wright's (1999, pp. 14-15) directives on how to study historical state's systems in a comparative fashion, including the suggestion to interpret European history as a succession of hegemonies not multiple independences as mainstream IR theory would have it. Wright suggests that comparing relations of authority and legitimacy across state's systems may be particularly fruitful for a theory of the state's system. He also takes Eeren's (1999) classical view of a 'system of states' as the starting point, emphasising the role that shared ideas and cultural values can play in international order.^{@G} It has been argued that this framework was introduced by Eeren as a counterpoint to the supposedly 'atheoretical' and descriptive history practised by some of his contemporaries in the field of history reflected by the English school (Luan D Eittle, 1999, p. 14). As we have

^{@G} Needless to say, Eeren's 'romantic' (organic) approach does not share such cultural elements with the early modern (contractual) conception of 'state's system', such as the ones analysed in the previous chapter.

seen, most of Watson's narrative about Westphalia revolves around relations of authority and legitimacy, but, under the influence of Eder and Wright, the cultural and ideational side is also present. For example, he provides a lengthy account of the clash between the medieval framework of Christendom and the new humanistic ideas on statecraft first introduced by Italian Renaissance. The point is that Westphalia gave more strength and legitimacy to *de facto* humanistic practices and marked the end of the political application of the medieval worldview. The latter, in turn, had already been weakened by the fragmentation of Christendom in an internal revolution (or '*stasis*') in any case of a struggle that culminated in the thirty years war. By proceeding in this way, Watson incorporates some of the early English school directives on how to study history, as well as some of the influence of Eder's historiography, which stressed precisely the need for a theoretical framework of analysis.

In addition to the claims on the nature of historical inquiry applied to the formulation of international analysis, we may also highlight some of the deeper assumptions behind Watson's view of order and change in international society. (Most of the argument is phrased in terms of a cinematic analogy of a pendulum swinging across a continuous spectrum of categories of relations of authority and legitimacy.^{@A} This is a highly dynamic view of social interaction of a snapshot of international society will reveal a certain position on the spectrum, but a proper historical approach will account also for the way the pendulum moves. Another device that stresses this dynamic view is the metaphorical notion of 'tightening' and 'loosening', which Watson (1995, pp. 1-2) employs with reference to the cohesion of a given international society. It has to be said, though, that while the pendular movement has a clearly illustrative character, this latter figure of speech assumes the reality of social cohesion as a key element to political order. On the spectrum of authority and legitimacy, we are reminded that while useful, some of the categories of analysis should not be employed in too rigid a manner. They are ideal types, no more than broad categories which cover a considerable range of distinct individual phenomena (p. 1). On the other hand, the theme of expanding values, norms, rules and institutions from the 'core' of international society to its periphery is, perhaps, the closest we get to identifying this implicit assumption. The 'tightening' and 'loosening' political order is thus understood according to a dynamic view and seems to rely upon the notion that there is indeed a social mechanism of

^{@A} A useful metaphor for a theory of systems is the pendulum. Imagine our spectrum laid out in the form of an arc, with its midpoint at the bottom of the pendulum's swing, somewhere between hegemony and domination (Watson, 1995, p. 2).

cohesion in operation.

This leaves room for considerations on the kind of social ontology to which Watson subscribes. Given the prominent role played by cultural elements and ideas in this mechanism of cohesion, a radically materialistic outlook is out of question. However, it has to be said that the approach developed in *The Evolution of International Society* is more sensitive to certain material aspects than most of the early English school texts. Reflections on military advantage, as well as territorial and economic elements are not entirely absent. The aftermath of Westphalia is portrayed as a major shift in the balance of power as a result of the thirty years war, new modes of taxation and redistribution of jurisdictions (Watson, 1995, pp. 9A'5/5). In addition to this, Watson (pp. 5.G'55>) also looks at the dynamics of colonialism and what it represented for the overall pendular movement in the spectrum of authority. Nevertheless, in his view, material features, albeit relevant, are not sufficient to account for the operation of international society. The ideas and values behind the core institutions and processes are still the main focus of attention. In the language of social philosophy, the assemblage of international order is 'territorialised' or 'deterritorialisèd' in terms of both material and 'expressive' factors (see CeEanda, 5//=).

'Westphalia as the constitutional order of a modern system

Daniel Philpott's soft constructivism emphasises "qualitative change in international society via change in the 'constitutions' of world politics. A constitutional blueprint in this sense is defined as a set of norms, mutually agreed upon by polities who are members of the society, that define the holders of authority and their prerogatives (Philpott, 5//., p. 5). They are called 'foundational' for being 'authors of orders, denoting the polities who carry on war and business, and the contours of their powers' (Philpott, 1999, p. A=>). His 'constitutive grammar' operates via 'three faces' of authority. The first face stipulates legitimate membership for polities in international society. The second face determines the rules for attaining membership, drawing a line between members and non-members and prescribing principles for dealing with outsiders. The third face defines the prerogatives of members, their authority and possibility of legitimate action at the international level (Philpott, 1999, pp. F9'G/; see 5//., pp. A'5.). We may thus compare different international societies based on what they look like in terms of the constellation of the three faces. All constitutions contain all three faces; every constitution's depiction of them is its unique signature (Philpott,

5//., p..A). Change in world order is explained in terms of 'revolutions in sovereignty', meaning a major change in at least one of the three faces of authority (p.59). These are key categories to keep in mind when looking at the historical formation of international society. The rest of the theory tries to elucidate the mechanism whereby these revolutions emerge and, in turn, substantially alter the general pattern of world politics when they spread across the system.

Thilpott's account of constitutional revolutions depends on a broader theory relating 'ideas' to social change. His 'framework' of ideas' is supposed to explain two sides of a 'chain of events' in such moments of transition. Ideas that deeply contest the status quo will emerge and become influential. They will initially reshape individual identities, and also those of wide social swaths, in a process of mass conversion and diffusion. In both creation and diffusion of ideas, a certain degree of self-reflection on, instrumentalisation of, and socialisation into, the new ideas will occur. This is the earlier stage of a revolution. Several eternal 'circumstances' (historical, institutional or social) operate as mechanisms at this point, constraining and enabling this process (Thilpott, 5//., pp.6-18). The second stage involves social empowerment of ideas. In the political scene, this means the ability of believers in ideas to alter the costs and benefits facing those who are in a position to promote or hinder the policies that the ideas demand (p.19). In order to have an impact in the constitution and change of world order, new converts have to 'lobby' (as it were) the ruling groups of a polity, so that in turn it may act on such demands in interaction with other polities at the international level. It is this kind of interaction that eventually leads to constitutional change. The process may be interpreted in a 'rational-choice' fashion, in terms of converts altering costs and benefits for the ruling elites. Thus we ought to look in detail at 'couriers of ideas' in order to analyse how change happens. Key players like the general public, networks of committed believers and groups in government are the main couriers, but institutions as such, and even regular international and transnational interaction may play the role (pp.19-20). Because revolutions in the constitutional rules of world politics account for clear-cut historical change, the application of these categories on conversion and social power of ideas to how they altered the three faces of authority in concrete cases is a key procedure endorsed by this approach.

The impact of Westphalia in world politics, according to the theory, can be assessed, first of all, by an account of how exactly identity change is related to new interests, or demands for a new order of sovereign states. Thilpott (5//., p.5) sees in

the Protestant Reformation as a crucial spring of our state system and the root of modern international relations, due to its close connection to the generalised conversion process that triggered the chain of events leading to Westphalia. Theologians of the Reformation like Martin Luther and John Calvin theorised a relation of separate interaction between the institutional church and the civil magistrate, drawing a line between political and ecclesiastical authority (pp.555-6). These principles were later on incorporated in the official creeds and confessions of the Protestant churches and, when applied in practice, undermined the political authority of Rome and strengthened the independence of local princes (see HanCrunen, 5/./). There is a strong correlation between the depth and reach of conversion to Protestantism following a 'Reformation crisis' and the demands for a new international order going through the wars of religion of the 16th and 17th century.

In the four chief polities (or region of polities, in the case of Germany) that fought for Westphalia in the thirty years' war 1618-1648, Germany, the Netherlands, Sweden, and France had an interest in a system of sovereign states arose within a generation of its Reformation crisis. Although not integrally involved in the armed conflict for Westphalia, England, Denmark, and Transylvania, too, experienced a Reformation crisis and supported the anti-imperial powers diplomatically. Important, too, is the fact that none of the Catholic polities, the Catholic German principalities, Spain, Italy, or Poland, developed any interest at all in a system of sovereign states. They remained allies with the empire (Hilpott, 5/./, pp.556-7).

In the phase of conversion, ideas were diffused from their theological 'entrepreneurs' to the general public via sermons, catechetical teaching and pamphlets. Princes and other political leaders also developed a strong focus on Protestantism and its political implications. The new form of religion (and several practices derived from it) would often be stipulated by the recently converted leaders (Hilpott, 5/./, pp.556-7). Dissent was not just a matter of ideas anymore but it became a practical issue.

Hilpott (5/./, p.556) theorises the social power of Protestant views by explaining three distinct causal pathways with reference to couriers of ideas. Each pathway summarises specific forms of struggle between the old order and the new worldview in each main actor that subsequently fought in the thirty years' war and held a stake in Westphalia. The Reformation from below, illustrated by the case of most Protestant polities in the Holy Roman Empire, had a popular, bottom-up, character based on social power exercised through defiant religious practices, threats of rebellion, politics within cities and principalities, rulers asserting the power of their position, and broad participation in the armed forces as volunteers against Imperial and

3apal domination (3hilpott, 5///, p.5F5). &he Lpoliti"ue solutionM, in turn, argued for the secularisation of political rule and an interest in internal unity and e\$ternal independence, or a system of sovereign states predicated on the public role of statecraft. &his was the case, for e\$ample, in ,rance, where the Reformation led to a long period of civil war until a restricted degree of toleration was secured (at least temporarily) for the 3rotestant minorities. &he political compromise was reflected also in its anti'imperial foreign policy based on *raison d'KLtat*, a principle also reflected in ,rance's demands in 7estphalia (3hilpott, 5//., pp..59'FG). ,inally, LReformation from aboveM as illustrated by the case of weden traces the main impulse for social change to the monarchs of that time, who did their best to reform the country in a top'down fashion.

weden's struggle for sovereign statehood, connected to a preference in 7estphalia for a new order of independent states, was historically attached to the deep personal devotion of the monarch and his perceived duty to protect the 3rotestant polities in the #mpire (pp..FG'=). All three 'causal pathways' are connected to a change in the 'incentive structure' stimulating relevant actors, and defended against the counterfactual alternative that these actors would have sought a new states'system regardless of the new religious ideas.

lased on this account of conversion to new political views favouring 7estphalian order followed by an e\$planation of the social empowerment of the new ideas, 3hilpott (5//., pp.F/; F5'F) proceeds to a discussion of constitutional change. .=G@ is described as Lthe origin of modern international relationsM and Lthe most significant revolution in sovereignty to dateM because it established the constitutional legitimacy of a system of sovereign states, revising the three faces of authority. 7estphalia Lmade sovereign statehood a normM, changing the first face by undermining the hierarchy between 3ope, #mperor on the one hand and local princes of 2erman polities on the other. It also changed the second face by clarifying what it takes to be a member of international society (3hilpott, .999, pp.A@.'5). &he third face, in turn, was addressed in terms of a clarification of the diplomatic prerogatives of the 2erman princes in the treaties, such as non'intervention (3hilpott, .99=, pp.G5'F). &hese in the long run were crystallised in terms of the principles of non'intervention. &here are, however, a few caveats here. &he claim, thus, is not one of '9ournalistic' change overnight, li'e a headline on a newspaper catching the reader's attention to the novelty of the story. A system was not created out of thin air. Rather, 7estphalia Lconsolidated three hundred years of evolution toward such a system. In history, perfect fissures are

rare, but as historical faults go, Westphalia is about as clean as they come. The new constitution originally operated in the European states' system and was restricted to Christendom. Some practices and institutions of the Holy Roman Empire still remained in use, despite its decline. Finally, some traits of the Westphalian system (even the political theorisation of sovereignty itself) already existed before the settlement. Despite all caveats, the traditional narrative on Westphalia and its significance as a marker of systemic change is maintained.

Thompson's narrative implies a number of metatheoretical principles on the relation between theory and history in IR, the nature of order and change in social reality and also some critical dialogue with rival IR explanations. In history in IR, he clearly defends a kind of narrative whose events and landmarks are changes in constitutional authority (Thompson, 5//., p. GF). For this reason, empirical data related to Westphalia is selected on the basis of relevance to the explanation of constitutional change. In fact, Thompson deliberately redefines 'Westphalia' in a more flexible way, so that it may include the Peace of Münster signed in the same year between representatives of the Spanish monarchy and of the United Provinces, ending the long war of the 'Cutch revolt' (p. F.). Moreover, when the test of the Peace of Westphalia can be shown to have been ignored by some of the relevant actors, he is willing to relativise the relevance of the immediate outcome of the negotiations and the need to look beyond the mere test of a treaty (pp. 55-56). Thus, we are left under the impression that history amounts to a series of 'case studies' which provide evidence to test a pre-conceived theoretical framework. Data, then, are fit into this theoretical straightjacket.

In addition to this, Thompson's narrative technique is also subsumed to a tense confrontation between a couple of contradictory 'directions'. On the one hand, the claim that Westphalia inaugurates modernity in world politics is highlighted in bombastic terms (e.g. Thompson, 5///, pp. 5-6; 5//., pp. @; F/).^{@=} On the other hand, what he finds as evidence in the settlement itself makes him "qualify the main thesis and seek supportive evidence elsewhere (Thompson, 199=, p. GA). On occasion, he will even relativise chronology and factual accuracy, perhaps relying too much on secondary sources reinforcing the conventional wisdom that he wants to update and defend (Thompson, 5///, p. 5/@). For example, he correctly points out that the Peace of Augsburg (.AAA) gave German princes the power to regulate religious matters in their territories

^{@=} Thompson (5///, p. 5/9) admits that consummate fissures in history are rare, but declares that Westphalia is as clean as historical faults come, declaring modernity's victory after Westphalia.

but did not put an end to the internal unrest in the Empire. However, he conceives this as a step to the Peace of Augsburg, after which the princes' sovereignty over religion became accepted and respected and practiced (Shiltott, 1999, pp. 10-11). The problem is that, far from reinforcing the Augsburg principle of *cuius regio eius religio*, Westphalia settled the return of the religious 'map' to a 'normal' year, regardless of the local prince's religion in each territory. It worked for that time precisely because it rolled back some of the Augsburg stipulations (Trauman, 1999, pp. 10-11).

Not only are there implicit metatheoretical principles on the role of historical research in IR scholarship, there are also statements on the nature of the social world. As we can gather from his exposition, Shiltott's account of the 'framework' of ideas' puts more weight on agency than structure. Although the narrative focuses on the agency' side of identity change, in a 'classical' individualist fashion, on occasion there are allusions to the existence of social and material structures. Even if structures do not awe us, we still surmise that ideas do not arise out of nowhere, but always come out of some set of circumstances. In fact, besides being present in the 'origin', it is also in the 'diffusion' of ideas leading to Westphalia and then a new system. The identities and interests of states are shaped through their social environment (Shiltott, 1999, pp. 10-11; AF). The Reformation was not Westphalia's sole cause; long-term material trends contributed, too (1999, p. 10). Circumstances and environment are neither purely material nor purely ideational. Material structuralism cannot be ignored but it is not sufficient for a full account. Key for international theory is the explanation of how interests of polities change according to new identities, although interests may be shaped by ideal or material forces (1999, pp. 10-11). In either case, the assumption is that this kind of social theory demands a reflective conception of human beings, rather than what Shiltott (pp. 10-11) perceives to be a more deterministic view advanced by structural theories. However, this conception is followed only to a certain extent, for in the description of the 'social power' of ideas, the cost-benefit analogy with rational choice theory is made under the assumption that new identities have already taken form. Hence the sharp distinction in the Westphalia narrative between 'conversion' to new ideas and their 'empowerment'. Identifying such presuppositions makes it easier to understand the framework of social change on which the theory draws.

On the nature of change, there is a number of scattered points in the approach. Shiltott's attempt to reconcile the 'conventional wisdom' on Westphalia with his own research findings translates into further tension at this level. On the one hand, his view

of change accommodates to the 'conventional wisdom' that Westphalia suddenly altered the structure of world politics as its founding moment (Hilpott, 1991, p.566). On the other, a list of caveats about the nature of the evolution of international society is provided. Westphalia was not an instant metamorphosis. Change spread across the system from Europe to other places gradually (Hilpott, 1991, pp.5/9; 5.F). Another tension on this issue has to do with Hilpott's view of 'cause'. He often describes ideas and structures as 'mechanisms' which constrain and enable certain outcomes. Ideas are also seen as constitutive mechanisms with reference to identities in 'conversion' (Hilpott, 1991, pp.5=; G=; >/). Ultimately, however, the historical narrative on Westphalia follows a more traditional 'sequential' pattern, a 'billiard ball collision' or 'domino effect' view of causation.

Iconoclastic propositions challenge the legitimacy of an existing international order, a contradiction that erupts in the volcano of the wars, the riots, the protests, the politics that then brings in the new order. This, through a *typical chain of events*. The ideas convert hearers; these converts amass their ranks; they then demand new international orders; they protest and lobby and rebel to bring about these orders; there emerges a social dissonance between the iconoclasm and the existing order; a new order results (Hilpott, 1991, p.6, emphasis added).

One event leads to another, until social change is explained and connected to the peace settlement of 1648. By following this conception of causation often recommended by the mainstream IR canon, Hilpott makes a further move that allows him to travel between (constitutive) constructivism and (sequential) mainstream theory.^{@>}

Finally, also connected to Hilpott's critical dialogue with other approaches is the defence of his view against possible defeaters of his thesis. In order to test the alternative claim that ideas did *not* lead to those changes in Westphalia, he provides a set of metatheoretical directives on counterfactuals. A valid claim in this vein designs an alternative world where the alleged cause is extracted yet other events and conditions remain intact. Another condition is *contenability*, or a theory of how, in this plausible alternative world, the remaining events and conditions might have brought about the same result. Hilpott finds such a world in which systemic change occurs apart from ideas in structural materialist accounts. While they meet this metatheoretical standard of plausibility, the actual test involves a more refined process (1991, pp.5/5). After each claim about 'conversion' to new ideas and their 'social power', a hard materialist account and an intermediate account are also tested

^{@>} Contrast with Geschle's view of causation, analysed below.

against Shillpott's own soft constructivism. For this reason, he insists in finding evidence that is also acceptable to these alternative approaches, speaking in terms of strength of 'correlations' and test of alternative hypothesis ' LBo Reformation, no WestphaliaM (Shillpott, 5//., pp.AA>; A@'=>; 9@'./G). &his helps us understand why more time is spent with the analysis of the behavioural implications of the new ideas than with a content analysis of the ideas themselves (compare Shillpott, 5//., pp../G'../ against most of the boo!).

'estphalia as consolidation of a pre3modern order

Cespite its title, Ienno &esch!e's *The Myth of 4*G*, emphasises primarily medieval and early modern world politics rather than the Westphalian settlement itself. &he 'myth' to which it refers is the classic story portraying L.=G@ as the origin of modern international relationsM. It is widespread in IR, I3& and I3#, whose students Lare united in invo!ing the Westphalian states'system as the benchmar! for measuring the present'day structure of world politicsM (&esch!e, 5//F, pp..'5). &his narrative Lhas given the discipline of IR a sense of theoretical direction, thematic unity, and historical legitimacyM. &esch!e (5//F, p.F), in turn, would argue that L.=G@, far from signalling a brea!through to modern inter'state relations, was the culmination of the epoch of absolutist state formation; it mar!ed the recognition and regulation of the international 0 or, to be more precise, inter'dynastic 0 relations of absolutist, dynastic politiesM. &he rationale for loo!ing at longer periods of time before and after Westphalia in order to assess the role .=G@ played (if any) in terms of social change has to do with some of the author's theoretical and metatheoretical choices. 7ritten under the influence of (ar\$ist historical sociology, *The Myth of 4*G*, highlights the processual development of relations of production, use of coercive power and control over land and 'class struggle'. Along the way, the boo! ma!es a number of metatheoretical points in connection with the general IR literature on Westphalia and the relation between theory and history.

<ey to &esch!e's (5//F, p.>) argument is what he calls a 'theory of social property relations'. uch theory has a clear materialist basis as a 'generative structure' which constrains and enables the reproduction and transformations of the system. &he e\$planation emphasises practices Lwhose construction, destruction, and reconstruction mediates humanity's metabolism with nature, while centrally implicating politics and geopolitics% social property relationsM. &he !ernel of this approach stipulates that Lthe constitution, operation, and transformation of geopolitical orders are predicated on the

changing identities of their constitutive units, which, in turn, are defined by configurations of social property relations. Geopolitical orders are time-bound balances of social forces expressed in terms of institutions that stipulate class-specific, and therefore antagonistic, rules of reproduction (p. 7). When this wider framework is applied to world politics, the operation of each particular system in history must be seen as a function of the prevailing property relations in its units (Eche, 1990, p. 5). In each type of system a definitive set of property relations generates specific geopolitical authority relations governing and setting the limits to inter-actor rationalities (Eche, 1990, pp. 5-6). This is what Eche (p. 7) calls a 'generative structure' - the element that explains institutional and behavioural differences. Despite this, there is no deterministic or 'closed' connection between property regimes and class-related strategies of reproduction (pp. 9-10). For one thing, change would not be possible, but we know that in a general crisis the system is very likely to be contested at the deeper level. That is, we see a 'transformative logic' accounting for qualitative change and systemic shift. Another reason to avoid determinism is that we need to acknowledge the lack of uniformity across units even within the same system, what Eche (1990a, p. 11) calls 'socially uneven and geopolitically combined real development of the regionally differentiated course of history'. When we study the broader level of analysis of world politics, we come to the realisation that the developmental potential of regionally differentiated sets of property regimes generates inter-regional unevenness. This leads to international pressures that spark socio-political crises in 'backward' polities, while the political and geopolitical responses to internationally induced crises react back on the international scene (Eche, 1990b, p. 11). At any rate, both reproduction and contestation emerge as active and conscious processes within the given constraints of the social structure (Eche, 1990, p. 7; see p. 9).

From this perspective, then, modern world politics would be intrinsically connected to a major transformation of the economic structure in its mode of production. The Marxist argument on the crucial changes brought about by the transition from feudalism to capitalism is well-known. Eche (1990, pp. 5-6; see 1990a), however, wants to account for the historical formation of our states/system highlighting the differentiated way in which the shift occurred. Modern international relations have at least two elements which were absent in the medieval world, namely, an inside/outside demarcation and an internal distinction between the political and the economic emerging within a capitalist economy. Feudal geopolitical order gave

way to certain institutions and patterns of class conflict between lords and peasants. Feudalism involved military coercion and appropriation, hence the lack of differentiation between the political and the economic aspects (see Emswiler, 1990, pp. 11-15). As for the absence of a clear-cut distinction between polities, since individual lords could simultaneously hold land in different kingdoms (O), feudal territoriality remained heterogeneous, shifting and unbounded (Sussman, p. 11). If we look at the historical details of the transition, contrary to most IR theories, we should come to the conclusion that Westphalia does not mark the transition in any of these key elements. Much of the rise of the political pluriverse of the states' system was the result of a long history of class conflicts over the sources of income that began in the tenth and intensified in the fourteenth and seventeenth centuries (Sussman, p. 11). As for the sharp differentiation between the economic and political aspects, we need to look at the emergence of capitalism after Westphalia, beginning with England and then spreading across the system (Sussman, pp. 11-12). Both transitions, therefore, were not triggered by the settlement of 1648. In light of the theory of social property relations, Westphalia was clearly not the starting point of modern international relations.

In light of this approach, it is possible to locate a transition period or, rather, a system on its own between the medieval and the modern world. Even with the emergence of a system of several (imperfectly) centralised states, European political practice still reflected the order constituted by a pre-modern generative structure of property relations based on dynastic and personal principles (Emswiler, 1990, p. 11). Contrary to the usual narrative, the Westphalian settlement reinforced an absolutist and early modern type of sovereignty which did not entail a separation of public and private, of politics and economics, since sovereignty was personalised by the king as patrimonial property (Sussman, p. 11). Practices of internal coercion and taxation, government, territoriality, as well as inter-state conflict, reflected this dynastic character of the states' system even after Westphalia. Particularly relevant to the operation of this system were dynastic marriage policies for the purposes of accumulation, or institutions of diplomatic precedence according to rank and nobility of the monarch. Disputes and wars of succession connected to issues of inheritance were widespread. Balance-of-power policies followed a peculiar pattern of small-state elimination. In such features we see much continuity within the early modern system, before and after Westphalia (pp. 11-12). If anything, Westphalia was the culmination of the epoch of absolutist state formation; it marked the recognition and regulation of the international order, to be

more precise, inter'dynastic 0 relations of absolutist, dynastic politiesM (p.F).

&his framewor! of social property relations is predicated on a series of metatheoretical points on how theoretically informed historical en"uiry should be developed in IR. In passing, we also read about a series of principles that may be adopted in this process of applying dialectic materialism to the study of the formation of states'systems. +n the general relation between theory and history we find, first, a denial of the possibility of an overarching and timeless logic and direction in historical formation. LAny encounter between IR as a social science and history will have to start from the assumption that there is no universal covering law that e\$plains international conduct across the centuries, as there is no one general e\$planatory theory of historyM.)et, this should not lead us to Lintellectual abdication to contingencyM. Although history is Lnot teleologicalM, we may still ma!e sense of it by hindsight (&esch!e, 5//F, p.>). 7hen it comes to the IR discourse on 7estphalia, the use of .=G@ even as a helpful historical mar!er based on timeless and universal e\$planations poses a problem and re"uires critical reaction.

A line of tacit acceptance runs through the literature, passed down une\$amined from IR generation to IR generation. Cates cannot lie, and the more distant the dates, the less the willingness to uncover their social content, conte\$t, and significance. Lut periodi?ation is no innocent e\$ercise, no mere pedagogical and heuristic device to plant mar!ers in the uncharted flow of history. It entails *assumptions a out the duration and identity of specific epochs and geopolitical orders' as it implicates IR theories /ith respect to the ade\$uacy of criteria adduced to theori?e the continuity or discontinuity of international orders* (&esch!e, 5//F, p.5, emphasis added).

&his is not merely a neutral, intellectual, matter of idiosyncrasy. &here are deeper political and ideological implications behind worldview discourses of continuity and change in world politics. uch narratives are also instruments that crystallise the *status \$uo* and constrain emancipatory transformation.

International Relations is a *social* science. As a social science, it does not stand outside the daily reproduction of structures of domination and e\$ploitation (O). &he subsumption of international behaviour under one general covering law claiming ob9ectivity is as theoretically impoverished as it is intellectually debilitating. 3olitically, it is dangerous and all too often complicit with the aggressive policies of the hegemonic state. In some versions, it is scandalous (pp.5>F'G, original emphasis).

Awareness of historical singularity and the unpredictable character of historical change is, therefore, re"uired in the critical study of systemic transformations.

&his leads to the second point on theory and history. *The Myth of 4*G*, is

permeated by critical interaction with alternative IR theories and how they (mis)use historical evidence. Despite the recent trend to assess the issue of systemic transformations with reference to history as a surer guide to theory-building, &esch!e's (5//F, pp.5'=) verdict is that Lthe relative paucity of IR contributions to a historical sociology of international relations re"uires mobili?ing non'IR literatures on general historical developmentM. ,or this reason, he advocates building bridges between historical sociology, (ar\$ist political economy, 2eopolitics and IR (&esch!e, 5//=a). As for IR theory, besides the claim that some of the mainstream approaches have a shallow view of the possibilities of deep historical transformation, there is also an empirical criti"ue of their incomplete interpretation of the Westphalian system. Realism, for e\$ample, with its Lnaturali?ed great power rivalriesM and Luniversali?ed balance of powerM, cannot account for the peculiar forms of 'dynastic balancing' found in the early modern states'system. Constructivism, in turn, ignores the Lproperty'related social sources of identity'formationM and ignores the Le\$tra'normative conditionsM of that international society (5//5, p.F=). &his empirical confrontation is metatheoretically informed, as it relates to &esch!e's (e.g. 5//F, pp..G'.=; G5'GA) portrait of realism and neorealism as historically impoverished reifications of social structure, as well as constructivism's ontological assumptions about the primacy of ideational elements, opposing the author's materialist starting'point.

Besides these negative points on theory and history raised against rival approaches, there are also positive metatheoretical elements. ,irst, there is a deliberate attempt to indicate the compatibility between the theory of social property relations and materialist dialectics. Despite the re9ection of universal and ahistorical generalities, the lin! between general (ar\$ist social theory and the analysis of the operation and transformation of states'systems in world politics re"uires considerable width (&esch!e, .99@, p.F5=). 3articular states'systems must be theorised according to their peculiarities, but the metatheoretical 'bac!bone' connecting dialectical materialism to IR theory remains. &esch!e portrays 2eopolitics as the intermediate space re"uires to build this bridge. -e criticises classical geopolitical discourse with considerations on its political agenda, but finds useful elements in its emphasis on material space in connection to political power. (ar\$ist alternatives are then analysed and found wanting in their original form (5//=a, pp.F59'FF). ,ollowing his e\$PLICIT metatheoretical analysis of the infrastructure of both research programmes in 2eopolitics, he opens up space for the approach developed in *The Myth of 4*G.*, La theory of IR based on

dialectical meta-theoretical premises (1999, p. 10). Contrary to rival approaches in IR and I3#, this enables the author to portray continuity before and after Westphalia and to distinguish a separate early modern dynastic system before the emergence of capitalism at a later stage.

In addition to this, the framework behind the theory of social property relations is made explicit in metatheoretical terms. Deconstructing the 'myth of Westphalia' is not simply a matter of empirical accuracy or ideological challenge, but also a call for new ways of theorising. The starting point is to adopt a comparative approach, since we need to single out differences across states/systems in each historical period. Moreover, we are also called to examine the deep causes of structural variations of what Wessels calls 'causal in history'. This is actualised in terms of an account of social change in function of crises enabled by contradictory strategies of reproduction advanced by competing classes. Because of the way he puts this principle in use, it is safe to conclude that this look at structural causation is a counterpoint to the prevailing 'billiard ball' or 'chain of events' view of causal in history in IR.@@ Another point here is the deliberate attempt to identify and theorise the major agents and processes of systemic geopolitical transformations, a principle followed with consistency and clarity throughout his work (Wessels, 1999, pp. 10-11). After looking at agency and structure in the dynastic states' system, these elements are in turn linked to specific processes (1999, pp. 11-12). He moves finally to enable a short re-reading of the Westphalian settlement of Westphalia in light of the historical context. It comes as no surprise that the most relevant clauses of the treaties of Münster and Utrecht are 'demystified' only in brief, for the ideas contained therein are derivative from the dynastic geopolitical order that the settlement was designed to confirm. Wessels (1999, pp. 12-13) is following metatheoretical principles which affirm the precedence of material relations and only then moves on to the analysis of rules, norms and ideas as a function of these relations.

Discussion

It is quite clear that the narratives provided by Watson, Hilpott and Wessels on the historical role played by the Peace of Westphalia in shaping the early modern states' system are intrinsically connected to their respective theoretical frameworks. As discussed, these theories and the stories told about Westphalia which they enable, regardless

@@ I have already referred to differing conceptions of causation in Hilpott's case. Another comparison between both approaches to causation can be found in this thesis (Chapter 3).

of their level of sophistication, are shaped to a large extent by metatheoretical views on a number of issues. In Table 5 (below) I have summarised the main points of comparison and contrast along theoretical, historical and metatheoretical lines. Disposing these elements side by side helps deal with the discrepancy between these accounts of Westphalia. Even much of the IR literature on Westphalia, the three authors make an effort to go beyond description and to provide a theoretically informed perspective on the event. The differences are not merely a matter of factual inaccuracy or discrepancy (though on occasion this may well be the case), but mainly a function of this choice, since the evidence is selectively reported according to a selection of relevant data in light of what the theory dictates (see 'Account'). These theoretical arguments and their applications to the case, however, assume different ways of bridging between philosophical issues of ontology, the nature of 'social order' and 'change' and the nature of historical formation and how it should be studied ('-istory in IR'). In the texts analysed here, the theories are also shaped by critical metatheoretical interaction with rival approaches ('Dialogue and critique'), although this is less the case for Watson's text.

Table 5 Three IR accounts of the Peace of Westphalia

	Teschke	Philpott	Watson
Theory (i.e. theoretical and conceptual framework)	Processes in world politics depend on identity of units in states' system. Identity understood as structure of social property relations. Processes understood as dialectical tension between classes.	Processes in world politics depend on constitution of international society. Constitution understood as rules, norms and institutions.	Processes in world politics depend on relations of authority and notions of legitimacy. These relations are constrained and enabled by shared cultural elements and degree of interaction.
Account (i.e. application of theory to the case of Westphalia)	Westphalia formalised a pre-modern, dynastic, states' system already in operation. It did not alter the fundamental nature of world politics. The modern states' system was, instead, triggered by capitalism later. The emergence of both dynastic and modern societies was geographically uneven.	Westphalia was a revolution in the constitution of international society. It set forth the parameters of modern world politics. It defined its fundamental principle of sovereignty, its prerogatives and membership criteria. Another defining moment was decolonisation after World War Two.	Westphalia formalised an international society of sovereign states based on anti-hegemony. Westphalian practices and institutions are embryonic of contemporary world politics, but went through incremental change over the centuries. They expanded from 'core' powers to colonies.

History in IR (i.e. metatheoretical points on how one should theorise world politics with use of historical cases)	-istorical states'systems understood in their own terms. (ain focus should be on peculiarities of each system related to social property relations defining it.	-istorical evidence shaped as 'case study' of the phenomenon of 'revolution in sovereignty'. Also used as 'test' of alternative theories in a counterfactual fashion.	-istorical states'systems understood in their own terms with a 'teleological' view to current system. &heory obtained from close empirical observation.
Nature of social reality (i.e. metatheoretical points on ontology of normal interaction)	3rimacy of material factors. &hese, in turn, define social classes and institutional framework's. Ideas not discarded. 3rimacy of structure. (entions agency.	3rimacy of ideational factors. Identities are shaped by ideas. (aterial factors not discarded. 3rimacy of agency. (entions structure.	'+rganic' cultural conception of international society. (aterial factors also relevant, but shaped by relations of authority:legitimacy.
Nature of social change (i.e. metatheoretical points on ontology of "ualitative change)	Institutional and practical change triggered by systemic crisis emerging from contradictory logics of reproduction.	Institutional and practical change triggered by revolutions in ideas if they spread across the system.	Institutional and practical change triggered by combination of ideas and change in relations of authority:legitimacy.
Dialogue and critique (i.e. metatheoretical bridges and antitheses)	Cialogue with historical sociology and (ar\$ist political economy. harpely critical of ahistorical features of mainstream IR. 6ritical of idealistic assumptions in rival views.	Cialogue with 6onstructivism. 3articularly critical of static IR theories and materialist IR theories. *ses them as 'counterfactual' worlds to test the theory.	Cialogue with early #nglish chool and theory'oriented historians. Implicit contestation of 'anarchophile' IR theories. Bot much discussion of rival approaches.

ource% +wn elaboration.

&he metatheoretical elements discussed in the wor!s of 7atson, 3hilpott and &esch!e also reflect the typology of metatheoretical discourse I have been adopting so far. &he points contained in &able =.. can be transposed to &able =.5 (below), represented by letters in the intellectual:conte\$tual and internal:e\$ternal *spectra*.

Ta le *!(5 Selected theory0history discussions in the typology of metatheory

	IB&#EE#6&*AE	6+B&#V&*AE
IB&#RBAE	<i>B</i>	<i>J</i>
		$\leq I_1$
		$\leq I_4$
#V&#RBAE		"P ₄ "P ₁
		T ₄ T ₁

ource% +wn elaboration.

&he theme of using historical data with reference to IR theory (letter *J*) is Internal to the discipline and related to 6onte\$tual factors (empirical evidence). Cialogue with other IR

theories, whether critical or not, falls similarly into the Internal category, but its focus falls more into the Intellectual side (letter *B*). The bridge between a broader view of historical development, how we should study it and IR theory is more to the centre (letter *I*). The metatheoretical discourse dealing with these issues is itself intellectually focused, for its main function is to apply a more general philosophy of history (*I*, external) to IR theory (*I*, internal). The origin of this bridge (philosophy of history) is external to the discipline of IR, whereas the destination (IR theory) obviously lies within the disciplinary scope. The other bridge we can see here is that of metatheoretical discourse on the nature of social order and the nature of social change. These two separate elements in Table 1 are joined in Table 2, represented by letter *C* also on the Intellectual side. The bridge is between social philosophy (*C*, especially ontology), external to the discipline, and IR theory (*C*). These are the points on which all of the authors have something to say. Individually, there are also a few elements that we could add to the table. Watson and Hilpott both make statements on early modern political theory and how it had an impact on statecraft and international dynamics (see "P"). This kind of argument relies on the study of theory (in this case, political thought) as a tool of the study of international society. Hence, it has an Intellectual focus (but external to the discipline, since it did not exist) moving towards the External side (social interaction). The same applies to Schlesinger's critique of 'ahistorical' IR theory for emphasising lack of change and, hence, preventing political emancipation today by implication (see *T*). The difference, of course, is that he has in mind the impact of current IR (Internal) discourses on Westphalia in the current system. We see, therefore, that the literature analysed here further illustrates some of the general contributions of metatheory outlined by my typology.

Watson, Hilpott and Schlesinger imply metatheoretical assumptions shaping their scholarship. Sometimes they are even clear about them. At any rate, their main goal is to study historical world politics empirically in light of their theoretical approaches. The position of the metatheoretical 'drive' behind their narratives, however, reflects a slightly different set of concerns to most of them matched by the recent critical literature on Westphalia narratives in IR. The question raised by such contributions is, what is at stake in these claims about Westphalia? It is unsurprising (but perhaps not stressed enough so far) is the extent of engagement with metatheory emanating from these works (see Table 3.F).

Table 5 Roles of the metatheoretical study of discourses on "testphalia

	Internal	Contextual
Internal	R	S, S _i
Contextual		Q P _i P, P _A

Source: own elaboration.

Similar to Rescher's critique, the literature questioning the traditional narrative on testphalia makes three key metatheoretical points. First, there is the issue of rigorous scholarship when it comes to an understanding of large-scale change in the history of international relations. This is Internal, being specific to IR research and mostly Intellectual, although the empirical reference drives it toward the Contextual side (letter P). Secondly, there is a protest against turning 'testphalia' the historical event into a yardstick or ideal-type to measure any occurrence of systemic change today. Similarly, this is an Internal and Intellectual metatheoretical statement, but approaching a Contextual element, as it refers to current social change (letter Q). Finally, the traditional discourse on 'G' has been accused of producing deleterious practical implications, both in IR scholarship and political practice outside academia. The accusation is against IR research and, therefore, Internal, but the impact is Intellectual in the case of its constraints on further IR research (letter R), and 'Contextual' when it comes to political practice (letter S).

The first complaint, the denunciation of the fact that sloppy historical scholarship has constrained and enabled specific moves in IR theory, is notorious, for example, in the critiques of John Gerard Ruggie (1990, pp. 19) and Stephen Krasner (1996) against the mainstream of the discipline. For both, a more factually accurate reading of testphalia would disconfirm the impression of 'normalcy' of our international system projected back to early modern Europe. Challenging this assumption, in turn, leads to a demand for significant theoretical readjustment in the mainstream. Speaking less about re-theorisation and more about the use of history by IR scholarship in general, others have also stressed the effects of poor scholarship in the discipline. Stéphane Lévesque (2006, p. 9), for one, concludes that the Lorthodoxy according to which 'G' can be credited for the birth of the modern state system is unsupported by historical facts, and is hence a myth. Jonathan Javerock (2005,

p.5.), for another, briefly summarises the mainstream narrative and then remarks that 'Anyone who has taken an introductory history course will realise that in practice history is not so neat. And critics' despair, the contributions of the revisionist literature on Westphalia have barely been incorporated and the mainstream of the discipline has failed to enter into any kind of dialogue with them, reinforcing the narrative's mythical character (Barvalho et al., 5/..., p.25). In other words, IR's credibility is 'questioned due to its poor treatment of historical evidence and, importantly, the philosophy of history it seems to assume. (Moreover, since Westphalia is a pivotal case used in support of a number of theoretical formulations, the rigour of IR theory is also 'questioned by implication.

Emerging from this critique is a further point, connected to the issue of large-scale change. On several occasions, the 'orthodoxy' about Westphalia is taken for granted and promoted to the status of a 'shorthand' expression for general political phenomena (e.g. 'sovereignty') in the term 'Westphalian'. This, in turn, is employed as a 'rule' against which contemporary world politics is checked in order to assess the degree of deviation of changes we can see today from the Westphalian 'norm'. Westphalia, in this case, is turned into an all-encompassing construct in its own right (Schmidt, 5/..., p.15). (Measuring current world politics and its effects against such yardstick becomes problematic when the baseline for comparison is itself problematic, especially considering that its use as a starting point for investigations of change may lead scholars to exaggerate the magnitude of recent developments or to assume a linear progression from some Westphalian configuration toward some 'post-Westphalian' state of affairs (Schmidt, 5/..., pp.16-17). Far from being a merely neutral construct, Westphalia is also evoked when current changes are addressed also from a normative perspective. And the extent that the historical example is adduced as a model for some sort of institutional arrangement either to be emulated or to be dropped, a lot hinges on the historical claims made about the peace of Westphalia (Trauman, 5/..., p.25). At stake here is the fact that this slanted history becomes a perspective and an interpretative technique that distorts our understanding of contemporary issues (Ayaoglu, 5/..., p.9). In short, a key criticism emerging from current metatheoretical discussion on IR theorising of Westphalia is that it has been extrapolated to interpret current events and transformation in a deeply contested way.

The fact, however, is that discourses of change and continuity in global politics can have a highly significant practical import. A final metatheoretical critique of the

traditional IR narrative is that it favours certain principles and excludes others. In scholarship as such, this is intensified by the conservative attitude of keeping the 'conventional wisdom' on .=G@ in the te\$tbody's, sometimes with a caveat indicating the e\$istence of alternative views without incorporating them. As a result, Lthe myths of yesteryear are perpetuated in the minds of generations of students as they in turn embar! upon their 9ourneys into the world of IRM (6arvalho et al., 5/.., p.>F=). uch perpetuation often has a constraining effect on the possibilities of discourse within the discipline%

X&Ythese historical insights (O) represent radically different perspectives or 'ta!es' on the discipline that necessarily confront the normalised and settled sub9ect matter. &he myths have had a tremendous function in disciplining our thin!ing about fundamental issues in international politics, 'normalising' it as common sense and providing the parameters or outer boundaries within which the disciplinary field is contained or homesteaded (6arvalho et al., 5/.., p.>A=).@9

l esides, the discursive and practical 'closure' enabled by the uncritical assimilation of the traditional view on .=G@ e\$tends to the e\$tra'academic realm. As Rob 7al!er (5//=, pp.A='9; =A'9) suggests, when it comes to political practice, the naturalisation of this 'truth'regime' has a conservative bias, in that it crystallises the *status \$uo* (the '7estphalian system') and discourages its confrontation by alternative discourses. &he LmythM, says leaulac (5//G, p..@=; see pp.5..'5.5), Lhas carried e\$traordinary power within the shared consciousness of society, including international society, and continues to impact discourses on contemporary issues on the international planeM.^{9/}, or e\$ample, the assumption that .=G@ satisfactorily settled the issue of 'difference' in

@9 &he main criti"ue is that the traditional narrative reflects a #urocentric view of the e\$pan\$ion of international society. &here are a few additional points which are more specific. &a!e #dward <eene's (5//5, pp.55'59; F@) criti"ue as an illustration. Cdescribing the #nglish chool approach to international society in connection to the settlement of .=G@ (including 7atson's narrative), he calls our attention to its intellectual ancestry in -eeren and other .9th'century conservative historians. &hey were motivated by the agenda of normalising an 'anti'hegemonic' view of the states'system against Bapoleonic revisionism. According to <eene (5//5, pp.F@'F9), Lwe have an historical narrative of the development of political and legal order in the modern world that is blin!ered in its outloo!, not by any real consideration of what is significant to us today, but rather in accordance with the needs of nineteenth'century reactionariesM. Adding to this, <ayaoglu (5/./, pp..9>'5/G) shows that our current view of political and legal order also reflects the rise of legal positivism of that time. -aving dismissed natural law theory, .9th'century international 9urists portrayed 7estphalia as the constitutional beginning of a society of states based on positive law. Assuming the classical story on .=G@ means reproducing this view on the foundation of international law. In sum, an uncritical absorption of the mainstream narrative brings some political 'baggage' with it, and shapes the discipline in function of this bac!ground. L+ften without reali?ing it, the numerous scholars today who use concepts li!e the '7estphalian system' (O) are therefore committing themselves to a peculiarly narrow and twisted perspective on order in modern world politicsM (<eene, 5//5, p.F@). 9/ &he point is made with reference to "quotes from documents written in the conte\$t of the International Criminal 6ourt.

domestic and international politics is a disincentive to think about new solutions 'outside the Westphalian box'. The problem is magnified where our celebration of modernity continues to blind us to the inadequacy of the guidance bequeathed to us by Westphalia (Inayatullah & Dillane, 2006, p. 67). Whether we agree or not with the diagnosis is beside the point. The relevance of the metatheoretical analysis of IR discourses on the legacy of Westphalia has been defended by these contributions in wider terms than mere 'scholarship about scholarship'.

Final remarks

The case of IR scholarship on systemic change connected to Westphalia is a very appropriate illustration of the way metatheoretical 'drives' operate with reference to the intersection between history, IR theory and social theory. Watson's English school approach reinforces the 'classic' view of the discipline on Westphalia as a 'paradigm shift' in the state system. The focus lies specifically on issues of legitimacy and authority. Europe's adoption of early modern, humanistic, political principles of statecraft and subsequent pattern of anti-hegemonic alliance formation are key evidences supporting the argument. Hilpott's 'soft constructivism' understands, first, as a revolution in sovereignty and the constitutional ideas in international society.

Secondly, it explains large-scale change in terms of how couriers of new ideas altered the cost-benefit structure leading to an empowerment of their demands for a 'Westphalian' system. Aware of the criticism against traditional views, he provides a more nuanced (and somewhat contradictory) view of change. It was gradual, but at the same time we are left under the impression that something suddenly happened in Westphalia. Eschle, on the other hand, dismisses the predominant account as historically inaccurate, theoretically flawed and politically harmful. His view is that, if anything, Westphalia consolidated a pre-modern dialectic order of social property relations, one based on dynasticism. The argument is predominantly based on material considerations. Key to modernity was the rise of capitalism, which happened much later.

As I have indicated, there is factual disagreement (even on what the term 'Westphalia' means), theoretical tension but also metatheoretical differences between these authors. Not only that, some of the nuances in each narrative can be ascribed to the metatheoretical infrastructure of the arguments. In their own way, the three contributions make reference to principles behind the use of historical evidence in IR

scholarship, an implicit bridge between social philosophy and IR theory of order:change, and critical interaction with alternative views in the discipline. On top of these shared issues, there are particular metatheoretical points in each work that I have stressed in my exposition. Having identified these elements with an in-depth analysis, I then proceeded to their classification in terms of the typology introduced earlier in this thesis. The results of my analysis fit neatly into the typology. In an additional procedure, I have also interpreted some of the recent IR literature claims to be the role of addressing discourses on Westphalia in metatheoretical terms. There are critical concerns about the quality of historical scholarship in the discipline, the use of the classical view of Westphalia as a framework for interpreting current events and the impact of this 'orthodoxy' on IR research and political practice. Those remarks were also classified in terms of my typology. Their availability adds to my argument on the role of metatheory at this time, with a specific reference to the study of theories on Westphalia and social change.

As in the previous chapter, I confirm once again that, across three different approaches to the same object, there is indeed a remarkable impact of metatheoretical elements on research. Implicit as it may be, metatheory plays a relevant role in 'driving' these three accounts of Westphalia to distinct positions. Notice, once again, that this is no deterministic relation. Metatheoretical discourse operates here as a constraining and enabling mechanism with an impact on the way systemic order, change and Westphalia are theorised and described. Unlike in the previous chapter, though, the texts selected for analysis belong to a well-formed discipline of IR. In this case, two differences emerge. First, there is a clearer line between 'internal' or disciplinary discourse and 'external' contributions of other disciplines. Watson, Chilpott and Beschle show awareness of the need to build intellectual 'bridges' between history, social theory and IR and make an effort to 'connect the dots'. Secondly, at least in Chilpott and Beschle the 'classic' disciplinary view of Westphalia is a background element, and so is the existence of self-identified rival IR theories that demand dialogue and critical response.

So far, I have tested my notions on the relation between metatheory and research in two concrete cases. The first case, presented in the previous chapter, establishes the metatheoretical influence on pre-disciplinary investigation. The case of IR theories on Westphalia, analysed in this chapter, does the same for disciplinary accounts. In each of these cases, I have offered an in-depth interpretation of core texts in each approach, elucidated implicit metatheoretical claims and discussed the role metatheory plays not

only in the cases themselves, but also in the disciplinary literature analysing the Westphalia narratives. Most theories assume a transition from 'hierarchy' to 'anarchy' actualised in 1648 and assume a generally anarchical configuration of the modern states' system. Alternative views bring up the notion of 'hierarchy' and apply it to the contemporary system. In the next and final chapter, I provide not only an interpretation of a variety of theories in IR and IJ on 'hierarchy' in the international system, but also apply an overarching metatheoretical framework to them, which purports to underline some of the 'Eurocentric' normative directives behind each approach. Chapter A offered a study of theory and metatheory connected to a conceptual normative debate in IJ. The present chapter contrasted IR theories in their interpretation of the impact of a historical event. The next will focus on a theoretical issue that has been attaining prominence in our discipline and in the field of IJ.

Chapter 5

Metatheory and a Theoretical Notion1

/ierarchy in 'orld Politics

Introduction

The centrality of the Peace of Westphalia in IR narratives about systemic change has to do, among other things, with the affinity shared by many theories on the postulation of the anarchical character of the international system. Unlike domestic political systems, so it goes, world politics is characterised by the absence of an overarching authority capable of imposing itself over the parts. State-centrism, another theoretical assumption dominating the field, reinforces IR's focus on anarchy. According to this theoretical assumption, states are the main actors of world politics and, at the interstate level, order and stability can only be provided in a suboptimal way considering the distinction between 'domestic' and 'international' political interaction. The more historically-minded IR theorist will, of course, point out as a caveat that this has only been the case since the 'system' changed from 'hierarchical' to an international 'anarchy'. In the discipline, the Westphalia narrative has little to do with the constitutional law of the early Roman Empire. Instead, it is a shared 'pool of facts' that enable the statement, from many theoretical perspectives, that our modern system began with the transformation of Europe from a 'hierarchical' society under imperial authority into an 'anarchical' system of sovereign states. In the previous chapter, we have seen that not all theories agree with this narrative, although most relevant approaches do. Part of what accounts for their distinct treatment of this 'pool of facts' has to do with different metatheoretical 'drives' constraining and enabling certain types of arguments. In this chapter, we move our analysis of metatheoretical mechanisms away from the historical and empirical type of research on Westphalia to the general conceptual issue of what 'hierarchy' means.

John Hobson's (2002) fresh work on the interpretation of Eurocentric views of 'hierarchy' in international thought will be eventually applied as an overarching framework of metatheoretical critique, but most of the chapter provides an 'internal', different sort of critical evaluation. It is asked, for each approach analysed here, whether

their theories of 'hierarchy' are consistent with the metatheoretical principles advocated by the authors themselves. The *structural-instrumentalist* view of neorealism, while focused on international anarchy, has something to say about 'hierarchy' and 'weaker states' and, besides, has been considerably influential in the formation of IR theorising on hierarchy. For this reason, I provide a brief discussion of this research programme with reference to asymmetry of states. An approach that is currently prominent in the field is the study of hierarchy, via *empiricist* theory, as a dyadic relation between a dominant country and its subordinates. Despite belonging to the mainstream, it differs from neorealism in many respects. While these two theories are well-known in IR and International Political Economy (IPE), the other two frameworks selected for analysis are not in the textbooks, but are by no means less worthy of due consideration. The first alternative approach is a *normative* theory of hierarchy influenced by both realist and liberal political philosophies. It is written from the perspective of the 'weaker states' and challenges mainstream 'state-centrism' in a number of ways. Importantly, this 'peripheral realist' approach presents itself in close dialogue with some of the IR and IPE literature, including a 'mercantilist' view of the wealth-power nexus. The second alternative approach, a *deductivist* theory of political economy, avoids any dialogue with the IR literature, although it employs international historical sociology to illustrate some of the historical points. This latter view defies 'mercantilism' and subscribes to 'liberalism' on the issue of the wealth-power nexus. Each of the four research programmes, therefore, has a distinctive metatheoretical 'drive'. Each of them are initially tried with the purpose to verify whether their substantive claims are consistent with their self-declared metatheoretical principles.

The next section presents a brief overview of the context behind the approaches selected for analysis here. Most theorists emphasising hierarchy at the international level engage in some sort of critical evaluation of neorealism, so this research programme is briefly examined first. After that, a discussion of mainstream theorising on dyadic hierarchy is provided, followed by an analysis of peripheral realism and, subsequently, the deductive theory in political economy. For an overarching metatheoretical evaluation, Hobson's framework is considered in its original application to neorealism, and then extended to the remaining empiricist, normative and deductivist theories. I ask, in passing, whether Hobson's own approach could be challenged based on this extension of its application. After that, I discuss some of the findings in this chapter about the role played by metatheory at several different levels. This will close

the third part of the thesis and the illustration cases, preparing us for the final conclusion.

A note on the content of the hierarchy literature

IR theory has historically inherited a liberal focus on the sovereign equality or 'freedom' of states, displaying what Iuan and Eittle (5///, p.GG/) have called 'anarchophilia', or the disposition to assume that anarchy at the international level is not only natural, but also a desirable thing. Kenneth Waltz's neorealist research programme, according to the author himself, begins with the 'theoretical fiat' that separates 'domestic' from 'international' political systems as a simplifying assumption (Waltz, .99/). Inter-state relations are perennially 'anarchical', contrasting with the 'hierarchical' organisation of domestic societies under the centralised authority of a government. Anarchy, in his view, accounts for the striking sameness in the quality of international life through the millennia (Waltz, .9>9, p.=). Since much IR theory has emerged either from liberal political views (e.g. Hull, .9>>), incremental improvement upon Waltz's neorealism (e.g. Bohane D Bye, .9>>), or critical dialogue whilst keeping an 'anarchophile' outlook (but see Wendt, .995), much talk of 'hierarchy' has been pushed to the background of the discipline.

Recent trends in IR and I3# bring the notion of 'hierarchy' back to the theoretical scene. This is the case at both theoretical and metatheoretical levels. At the theoretical level as we have seen in the previous chapter, Adam Watson and other key figures in the English school such as Martin Wight had already broken up with the predominant view that the international system is perennially anarchical. Instead, they theorised its functioning, development and change in terms of typical categories of authority and legitimacy allowing for variation in the symmetry of relations between the units. In I3#, dependency theory (see Cardoso, 5//>) and some of the approaches inspired by neo-Marxist analysis (e.g. Wallerstein, 5//G) postulated a differentiation between 'core' and 'semi-peripheral' states, ascribing to them distinct roles in the world economy, according to structural constraints and unit-level features. Moreover, the 'hegemonic stability' thesis was developed, allowing for a greater contribution of dominant states as providers of 'public goods' (such as order itself) in world politics (Zilpin, .9@>, p.@=). However, especially in IR theory, 'anarchophile' formulations advanced by neorealism, liberalism and constructivism still prevailed for most part of the time, often combined with Eurocentric and state-centric biases (Iuan D Eittle, 5///, pp.5/'). Other approaches

are analysed below. Alternative views, such as Carlos Escobar's 'realism of weak states', premised on 'hierarchy' from a 'peripheral' and 'anti-statist' perspective, were never that influential in mainstream debates. Despite this, hierarchy-based approaches have been gaining voice in the discipline (Lang, 2006). Few historical case studies, for example, challenge or qualify the assumption of sovereign equality embedded in balance-of-power models (Kaufman, Eittle, D Thohlforth, 2006; Thohlforth et al., 2006). David Easton's contribution to the debate emphasises bilateral hierarchical arrangements of defence and trade between the state and its subordinate allies. There are also political-economical theories which speak of international and hierarchical relations but do not pursue any explicit dialogue with the IR literature. Such is the case of Hans-Joachim Lauth's passages on the rise and decline of hegemonic liberal states.

At the metatheoretical level, the recovery of hierarchical notions is also a recent trend. In the past, the dichotomy between domestic 'hierarchy' and international 'anarchy' has been analysed from a disciplinary-historical perspective, focusing primarily on political theory. Some of the arguments predicate the distinction in terms of *similarities* between domestic and international political interaction and advance a normative agenda based on the inferred absence, at the interstate level, of mechanisms that provide domestic order. Others, instead, assume a qualitative *distinction* between domestic and international societies, leading to normative views that do not necessarily require a supra-national 'hierarchy' to provide order and stability (Kuganami, 1999). Recently, critics of colonialism in international practice and thought have undertaken research on the implicit Eurocentric assumptions in theories of world politics. In the case of the English school, for example (Glar, 1999, 2009), acknowledging the reality of hierarchy is said to be only the beginning of a critical approach from a postcolonial perspective. The assumption of hierarchy is usually turned into legitimization in its understanding as an 'institution' that promotes order in international society (Keene, 2005). Normatively speaking, the political consequences of such legitimization of hierarchical orders could be, and have been, catastrophic (Einiklater, 2007). John Hobson's study of Eurocentric assumptions in international theory derives from this context. One of the main findings in his analysis of theoretical material developed both before and within the discipline of IR is that, at least implicitly, hierarchy and 'differentiated sovereignty' have been a constant (not an exception) in international thought. After a critical discussion of whether Walt, Easton, Lauth and Escobar are coherent with their own metatheoretical views, I evaluate Hobson's metatheoretical

treatment of hierarchy in IR theories. I briefly recall his objections to Walt and extend the application of the framework to the new cases of East, Europe and Russia.

Hierarchy and great powers states in neorealism

Neorealism, originally developed in Walt's *Theory of International Politics*, is among the most influential research programmes in the discipline (Alinia et al., 5/..). In my analysis, therefore, I shall only emphasise those elements which are immediately relevant to this chapter.⁹ Walt's 'instrumentalist' view of theories has already been expounded in this thesis (see part one), but there are additional metatheoretical arguments worth mentioning. For him, a theory isolates one realm so that it may deal with it intellectually (Walt, 1997, p.5). In his view, 'systemic' or 'structural' theorising in IR (following his own neorealist approach) manages to abstract the realm of the 'international', separating it from 'domestic' politics. His simplifying move enables, in an instrumentalist fashion, the theoretical study of international politics (p.59). Systemic theories conceive of causes operating at the international level, in contrast with 'reductionist' theories, which concentrate on causes at the individual or national level. In this latter type of study the whole is understood by knowing the attributes and the interactions of its parts (Walt, 1999, p.10). Systemic theories explain processes by postulating an 'intervening' force between interacting units and the outcomes of their actions (p.10). Reductionist views, in turn, are criticised by the author for not abstracting the realm of international politics. As a consequence, they keep including *ad hoc* unit-level auxiliary hypotheses to account for unexpected behaviour (pp.59-60). The critique of *ad hoc* strategies is substantiated with a partial subscription to the 'methodology of scientific research programmes' of Imre Lakatos (see also Walt, 1997, 5/10, chapters 1-3). For Walt, therefore, systemic theorising is the way to proceed.

A system, according to the neorealist research programme, is a set of interacting units, including a system's level component of structure, which enables us to isolate the international realm as a whole more than just the sum of parts (1999, p.6). According to Walt (1997, p.67), structures shape and shove, operating as a constraining and disposing force, enabling us to explain and predict continuity within a system in a non-deterministic way (1999, p.10). We may define a 'structure' in

9. I have elsewhere provided a thorough exposition and critique of Walt's *Theory* in light of his later writings and the philosophy of science to which he subscribes (E. 2. Aire, 5/10a; 5/10b). (Metatheoretical studies on certain aspects of his work are widely available (e.g. Joseph, 5/10; <Eohane, 9/10; <Court, 5/10; <mpson, 5/10; <hweller, 99>; Hasue?, 99>).

terms of three elements% its 'ordering principle', the degree of 'functional differentiation' between the units in the system and the 'distribution of capabilities' across them. As we have seen, internationally, the 'ordering principle' is perennial anarchy. Because of the 'self-help' behaviour ensuing from the lack of a central authority capable of imposing itself over the parts, the relevant actors interact in a low degree of 'functional differentiation'% they do not specialise, but need to operate independently and provide for themselves (Kearsheimer, 1991, pp.175). This means that the two elements of the structure tend to remain unchanged. What is left to explain systemic difference between structures (and, consequently, patterns of interaction and stability) is the 'distribution of capabilities' (1991, pp.171-172). Walt's *Theory* is known for the prediction that a 'bipolar' distribution is more stable and lasting than a 'multipolar' one. The internal logic of the original framework is consistent with Walt's metatheoretical requirements of a systemic approach, as well as the metatheoretical criteria employed by the author to criticise his reductionist opponents (see Cini, 1992, pp.15-16).

'Hierarchy', in neorealism, is sharply distinguished from 'anarchy'. It applies primarily to situations internal to the states. It is in domestic, not international politics, that units of institutions and agencies stand vis-à-vis each other in relations of super- and subordination. By implication, in this setting political actors are formally differentiated according to the degrees of their authority, and their distinct functions are specified (Walt, 1991, p.172). The international system, conversely, follows a different logic.

Whatever elements of authority emerge internationally are barely once removed from the capability that provides the foundation for the appearance of those elements. Authority usually reduces to a particular expression of capability. In the absence of agents with system-wide authority, formal relations of super- and subordination fail to develop (p.172).

In the absence of a world government, states counterbalance the rise of a great power by forming alliances to contain it. Structure will reward some kinds of behavior which are compatible with its constraints (p.95), such as, for example, 'balancing' rather than 'bandwagoning', or joining the weaker side to stabilise the distribution of capabilities against a rising hegemon (p.15).⁹⁵ Other versions of neorealism will go much further in their exploration of the alliance-making practices available to great powers, but the

95 Balancing, or joining the weaker coalition, makes sense because after the distribution of capabilities has been balanced, the balancing state will have to face potential rivals in its own coalition. The weaker they are in this 'second round', the better.

structural claim remains at the general level ((earsheimer, 5/.). In the neorealist approach, therefore, there is a clear-cut distinction between anarchy (international politics) and hierarchy (the domestic system). &he structural elements derived from this distinction press states to avoid relying on any other unit. &hey tend to become very similar, and to prioritise survival as their highest goal.

But are states not different in any way? Of course they are. Walt's (.9>9, pp.5' F; 9F'=) point is that the theory accounts for Great Power behaviour at the systemic level, focusing on the most important states rather than each and every one of them%

It would be as ridiculous to construct a theory of international politics based on (Malaysia and Costa Rica as it would be to construct an economic theory of oligopolistic competition based on the minor firms in a sector of an economy. &he acts of all the states and of all the firms in a system are affected much more by the acts and the interactions of the major ones than of the minor ones.

Granted, if we isolate weaker states from Great Power intervention, then they will become the relevant units in their own system. But in this case the same theoretical principles would apply. &he choice to exclude weaker states as a focus of the theory is that they would not add much to a systemic and 'parsimonious' explanation. In a multipolar world, we only need to look at the Great Powers. In a bipolar world, the two superpowers. &his reasoning is definitely consistent with Walt's self-imposed metatheoretical requirement of a systemic framework!.

However, in order to extend the explanatory power of his balance-of-power thesis that structure will reward states that balance against a rising hegemon, the author introduces the notion of 'secondary states' and the example of the Peloponnesian War, when small units joined Sparta (weaker coalition leader), not Athens (stronger power). &he logic is the same% Secondary states, if they are free to choose, flock to the weaker sideM (Walt?, .9>9, p.5>). We know that, nevertheless, at least in the case of Athens, that decision to balance, rather than 'bandwagon', jeopardised its survival as a political unit (Mears, 5//G, p.FA). &his unnecessary *ad hoc* account of 'secondary states' that, in this case, fails even as an illustration is perhaps better understood in light of Walt's (.99F) later move to incorporate smaller states into the framework in order to make predictions after the Cold War. One would expect a theory of 'unipolarity', since one of the two only relevant superpowers had disintegrated (Crawthorn, .99/; Wohlforth, .999). Instead, he argues that the * was also declining because of the 'balancing' behaviour of other states in the West. &hus, the asymmetry between the only

superpower and rising powers like Germany and Japan is downplayed (see also Eayne, 1996 in the same context). Besides, Walt's evaluation of these countries as rising powers breaks down their 'capabilities' in terms of 'sectors', emphasising economic and technological strength at the unit level, rather than providing a relational 'distributive' comparison between them and the * at the systemic level. Finally, the author also implies that there is a certain probability that the * will be a more 'benevolent' hegemon, balanced more slowly because of its lenience and reluctance to intervene, which is another unit-level feature added to the post-Cold War prediction (Walt, 1996, pp. 22-23).^{9F} One could do better than add such *ad hoc* auxiliary hypotheses to account for the end of the 'most stable' arrangement of bipolarity only a decade or so after Walt's *Theory* had been published (see also (Karsheimer, 1997). In neorealism, smaller states are either completely excluded, or partly included by the backdoor. This careless operation, facilitated by the instrumentally drawn sharp dichotomy between hierarchy/anarchy corresponding to domestic/international orders, has left many dissatisfied in the field. The asymmetry in the relations between the * and subordinate states after the Cold War is key to the 'dyadic' approach, one of the most relevant subsequent alternative views of hierarchy. In this case, still a mainstream formulation.

Mainstream theorising and dyadic hierarchy

Eaale begins his study on *International Hierarchy* with the claim that the concept in IR is alien, denied, and excluded because the most common assumption in the field is that of an international anarchy (5/9, p. i). The contention is that, while the system is anarchical, we can look at how groups of states within the system interact and still identify hierarchical patterns, where some states willingly subordinate themselves to another, but typically only for something in return. Hierarchy in this sense means the extent of the authority exercised by the ruler over the ruled and, most relevantly, in economic and military terms. Authority involves legitimacy rather than sheer coercion (Eaale, 5/9, pp. 2-3; 2.9). Eaale's study examines both sides of a dyadic relation of hierarchy in abstract, and applies the theoretical insights, via measurement indicators, to empirical cases of bilateral relations between the * and its subordinates (see also Eaale, 1997 for different illustrations). A number of propositions may be

^{9F} Elsewhere, Walt (1996) does a better job in connecting unit-level variables to the study of foreign policy rather than international politics, and subsequently differentiating both types of theory. Nevertheless, I maintain that their conflation is unhelpful in the other texts, because Walt's declared intention of parsimony and simplicity, condemning unit-level theories for their use of additional hypothesis when empirical evidence seems to contradict the predictions.

inferred from the nature of hierarchical authority, such as its variation in degree depending on the case, its reliance on consent and the problems of enforcement by, and self-restraint of, the ruler (Ea!e, 5//>, pp.A='=.). &he main point is that the hegemon provides to its subordinates some degree of Lpolitical orderM that is worth the relative loss of autonomy. &he terms of this e\$change are constantly negotiated, and the implication is that Lsovereignty is dividedM and varies in degree and issue'area across the system (Ea!e, 5//9, pp.5/'F/; A.'5). 6ontrary to 7alt?'s neorealist approach, dyadic hierarchy implies a different behaviour from self'help, resulting from a higher degree of specialisation, since the hegemon in this model provides some of the protection, predictability and stability (pp..='G).

,rom the introduction of such concepts as 'authority', 'legitimacy', 'sovereignty' and 'political order', we e\$pect a philosophical discussion, or a 'normative theory' in the sense described in 6hapter . of this thesis. +r, perhaps a 'social theory' interpreting how 'intersub9ective consent' creates hierarchical order between groups of states, including 'rules' and 'norms' of legitimacy, which 'constrain and enable' international interaction. Instead, the discussion is merely footnoted with political theory and developed as mainstream empirical theory (but see Ea!e, 5/./). Ea!e is more interested in measuring these things in terms of e\$ternal behaviour and material cost:benefit analysis (5//>, pp.='>>). -e wants to test hypotheses against rival theories, but only those approaches which would allow for the same !ind of evidence. (aybe this is the reason why he ignores much of the #nglish school wor! on hierarchy, already available for a long time in the discipline (e.g. 1ull D 7atson, .9@G; 7atson, .995; 7ight, .9>>). Instead, he opts for dialogue with theories li!e the 'hegemonic stability' thesis and 7alt?'s neorealism. An important disagreement with the 'hegemonic stability' model is that, for Ea!e, 'order' is not a global public good, for it can be (and is) provided in an e\$clusionary way. &his means that the hierarchical relation needs not be systemic, but rather restricted to the ruler and its subordinates with the same effect of stability (Ea!e, 5//9, pp.FA; =/). 7ith reference to neorealism, the point of contention is not only substantive (i.e. "uestioning the universality of self'help behaviour and functional similarity in the system), but also metatheoretical. ,rom a 7alt?ian perspective, Ea!e's theory would "ualify as 'reductionist' for focusing on units and their relations, rather than taling the general system as a starting'point.

&he author's preference for a mainstream empirical'theoretical approach seems to be at odds with some of his self'declared metatheoretical principles. &he construction

of the framework! and predictions is allegedly Lpremiered on a critical realist approach to theory and measurementM. 6ritical realism is defined as La postpositivist approach to science that presumes an independent reality e\$ists but our !nowledge of it will always be imperfectM (Ea!e, 5//9, p.=F; see note .). &he criteria for theory test, however, is provided, first, as merely replacing Lone flawed theory by a less'flawed theoryM and then with reference to the not'so'realist philosophy of science of Imre Ea!atos (Ea!e, 5//9, pp.\$ii; =A). 1ut a third approach to theory, defended in the introduction, is the one more consistently followed in the study. It e\$PLICITLY echoes 7alt?'s instrumentalist (and anti'realistS) view%

All theories are based on sets of simplifying assumptions that help render a comple\$ reality more easily understood. Assumptions are 9udged by the e\$planatory power of the theories they generate (O). 1ut it is important to recogni?e that these are not empirical descriptions of reality, but merely assumptions that we can accept or re9ect on their e\$planatory power (Ea!e, 5//9, pp.F'G).

#ven ne\$t to where 'critical realism' is declared to be the metatheoretical guidance behind the study, a mainstream:instrumentalist defence of operationalisation is assumed. L6onstructs without empirical indicatorsM, says Ea!e (p.95), Lare of little practical use and ultimately cannot be shown to be more useful in e\$plaining real world politics than the alternativesM. +perationalised variables, in his view, help us e\$plain by chec!ing whether statistical correlations are strong enough to postulate something li!e a 'law' (Lrieden D Ea!e, 5//A, pp..F>'@). &his is far from the alleged 'realist' position. ,or philosophical realists, scientific e\$planation is made not as regression analysis, but in terms of uncovering an underlying mechanism (1unge, 5//G). 7hile proper 'critical realists' would not tale issue with empirical research *per se*, this odd mi\$ture of instrumentalism, empiricism and the anti'realist stance on theoretical assumptions makes one wonder whether Ea!e is merely paying lip service to 'postpositivist critical realism' to avoid the negative labels associated to mainstream 'positivist' theorising.

#lsewhere, Ea!e (5/..) criticises philosophical and theoretical 'isms' for artificially crystallising metatheoretical and cross'theoretical boundaries. In an attempt to stimulate 'analytic eclecticism', he compares strict adherence to 'isms' to some sort of religious persecution of 'heretics'. 3erhaps this strange blend of 'isms' is simply his own attempt to be eclectic. 3erhaps we should understand his theory of hierarchy based on substance, and not the contradictory character of its metatheoretical 'drive'. ,air enough, but in this case, for the sa!e of clarity, I would have preferred him to simply reinforce

his *de facto* adhesion to mainstream theorising by declaring adhesion not to 'eclecticism', but to the 'sect' that actually *preaches* what he *does* 0 instrumentalist empiricism. &he reason is simply (but importantly) that adopting a position that is actually 'closed' under the guise of 'openness' can be more detrimental to a pluralist and 'eclectic' agenda, erasing real alternatives to the mainstream, than the coe\$istence of multiple approaches (Bau, 5/..).

Bevertheless, there is a number of ways in which Eale's metatheoretical commitments (especially the ones he actually follows) influence substantive research. &ale, for e\$ample, the obsession with re9ecting non'mainstream theorising when he proposes indicators Lbased on observable behaviors rather than institutions or intersub9ective understandingsM (Eale, 5//9, p.=>). In case one as!s why the theory is Linsufficiently socialM, he replies that an Lepistemological betM on the study of Lstrategic interactionsM was made, and that ultimately Lauthority rests on the largely material e\$change of order for compliance and legitimacyM. ,or this reason, he denies the relevance of Lideas and normsM (p.\$i). &his is "uite peculiar, given the connection between 'authority' and 'legitimacy' (based on the ideational notion of consent), and the definition of 'order' as Lthe protection of persons, property, and promisesM ' echoing -edley 1ull's (.9>>) norm'based approach (Eale, 5//9, p.59). &he 'epistemological bet' is based not on the research "uestion as such, but on pre'conceived metatheoretical standards of what ma!es good research. &he "uestion about hierarchy in world politics, authority, legitimacy, consent, etc. is as!ed, but the answer given is in terms of correlations between factors li!e Lpresence of military forces from dominant state, A, on the territory of the subordinate state, 1M or Lnumber of independent alliances possessed by 1, the potentially subordinate stateM (pp.=@'9). &he research is further restrained by this impulse in the selection of cases. &he * in its bilateral relations is the sole 'dominant state' being studied because Lsimilar data on these indicators are not available for other countriesM (p.=@). If pressed about the selection of these indicators, Eale responds that Lthere is no reason to prefer one (O) over the others. 7ithout being able to observe the inherently unobservable, we cannot !now which of these indicators is capturing more (or less) of the construct of hierarchyM (pp.>.'=). +ne wonders what could have happened if he had chosen to actually follow 'critical realism' (designed to deal precisely with this type of problem) as a metatheoretical framewor!, integrating empirical evidence to a more 'social' !ind of theorising. Cespite Eale's current popularity in the discipline as a theorist of hierarchy, some of the insights on the topic,

as well as criticisms of Waltz, have been in circulation for many years. This is particularly the case for the not-so-known approach of 'peripheral realism'.

Normative theory and hierarchy in peripheral realism

In spite of the title, #scudJ's *Foreign Policy Theory in Menem's Argentina* purports to introduce a mid-range explanatory and normative theory that applies to 'weaker states' in general. This 'peripheral realism' takes at face value the principle suggested by the Athenians to the Peloponnesians, that the weak suffer what they must (Thucydides 1.90, H.55, trans. R. Crawley). From this claim, he derives directives of political prudence in peripheral statecraft. *Foreign Policy Theory* looks at the interstate system from the perspective not of the powerful but of the powerless, while refuting alleged fallacies in mainstream theory that are more noticeable from the viewpoint of the periphery but that nevertheless affect the logical structures of mainstream international relations theories (scudJ, 1997, pp. 6-7). It is therefore clear, from the beginning, that the theory presupposes a hierarchical arrangement between the 'powerful' and the 'powerless'. Hierarchy, for scudJ, manifests itself in both *de facto* and *de jure* terms. An example is the veto power for some of the great powers in the UN Security Council. States are not formally equal; admittedly, they are even less equal on an informal basis, but it is an untruth to say that none is entitled to command and none is required to obey (p. 10). Rather than lament or oppose such hierarchical constellation from a peripheral perspective, the author takes it as his starting point and challenges, instead, 'eccentric' foreign policy that resists or ignores hierarchy. Adventures in the periphery, he claims, are unaffordable. This is particularly so if we break down the 'black box' of the state-as-unitary actor, defying a key assumption of mainstream IR theorising: when a small state attempts to behave like a great power, its government may benefit from militaristic nationalism, but the population suffers the consequences of international sanctions.

In fact, the 'questionable character of this 'state-centric' assumption is a central metatheoretical point reiterated by scudJ in every step of the argument. 'Scientific' theories at the core of the discipline are denounced for ignoring the normative presuppositions that they smuggle into supposedly 'neutral' explanations of power politics. The most notable issue is that mainstream international relations theory unintentionally takes sides with the state rather than with the citizen. Instead, scudJ defends a clear adhesion to 'political liberalism'. In his definition, the doctrine regards the individual human person's rights and interests as the foremost priority of the state's

activities and endeavors. It establishes the rights of the individual and interests of the citizen as the sole justification for the existence of the state (1997, p. 6). Peripheral foreign policy that confronts the hegemonic states risks the welfare of the citizens by exposing them to economic sanctions and other forms of punishment. It is mainly the ruling elites who benefit from this type of policy. Interestingly, the author does not extend 'liberalism' to the economic realm. Instead, he agrees with the 'mercantilist' interpretation of the wealth-power nexus and argues that the effects of mercantilist state-centrism designed to benefit the ruling elites can generate spillover into citizen welfare as an unintended consequence. This would be a benign kind of state-centric rationality, despite falling short of an ideal citizen-centric policy standard. The worse kind of state-centric logic is what we find in states which are obsessed with power: security and which threaten to impoverish the population, such as Argentina during the Condor II missile programme in cooperation with Iraq (p. 7). Being *realistic* in the periphery, therefore, often means ignoring some of mainstream realism's advice.

These are the normative elements behind the author's main point of metatheoretical contention against the mainstream. Let me further expound the contention itself. Critical of the instrumentalist assumption that states are unitary actors and of the metaphysical claim that 'states are people too' (Euomah, 5//9; Wendt, 5//G), #scudJ denounces such claims as an anthropomorphic fallacy or fiction of the group/person, leading to a tendency to think of state policy as if it were equivalent to an individual's decisions, and a political inclination to see the state as an end in itself, rather than being an instrument to defend the rights and welfare of its individual citizens (1997, pp. 5F'G). Anthropomorphism in this sense hinders our understanding of how foreign policy works and whom it benefits. This sort of argument is often the province of those known as 'critical theorists', drawing on the notion of a 'state/society complex' with distributive effects (Ashley, 1997, p. 60; Einik, 1997, p. 5). However, if #scudJ is right, a critical approach can also be employed by political liberals and shift the conversation into a more individualistic discourse that breaks down the state into different groups and structures. As a liberal criticising the distributive effects of the relation between ruling elites and the rest of the population, #scudJ's (1997, pp. F= >) dialogue with 'critical theory' appears to be more consistent with the principles of political liberalism than state-centric 'liberals' and 'neoliberals' in mainstream IR (but see van de -aar, 5//9, pp. 5A'A/ for alternative interpretations of liberalism in 13#). In fact, a convocation is issued to theorists, who are called to uncover and denounce the

discursive mechanism whereby 'anthropomorphic' principles (e.g. nationalism, militarism) are used to mobilise popular loyalty in support of grandiose and risky foreign policy in peripheral states (Hudson, 1997, p. 6A). &thus, even at the metatheoretical level, a political critique can be raised against the social consequences of applying mainstream state-centric theory in statecraft, and as a 'drive' to theorise IR and IR in different ways from a normative perspective.

Anthropomorphic 'fallacies' are a eloquent proof of the fact that philosophical assumptions are of necessity built into the very logic of international relations theory (Hudson, 1997, p. 6=), but there are additional aspects of metatheory explored by the author. &these relate, first, to his view of what makes statecraft 'rational' (in dialogue with Morgenthau) and, secondly, the role played by 'structure' (in dialogue with Walt?). In terms of rationality, Hudson follows Morgenthau's (1960, p. F) classical realist advice that we follow the forces inherent in human nature instead of working against them. In its peripheral formulation, however, realism attempts to further clarify the means and ends of statecraft in a 'centric' fashion, reflecting the earlier approach's state-centrism. &his is a theory-building tactic that yields a certain normative ideal type of foreign policy based on a conception of rationality according to which the ultimate ends must be moral, or good, and they must above all serve the people (Hudson, 1997, pp. 5'G). For this reason, Hudson defines his own approach as a political theory aggregating explanation and prescription. He is sceptical of its ability to fulfil expectations constructed for mainstream empirical theorising. Instead, he sides with Morgenthau's own conception of theory in relation to prudence and statecraft (1997, p. 7.).

&he other metatheoretical issue has to do with the notion of 'structure' in interaction with the neorealist research programme. Structuralism in Walt's conception has a very limited usefulness, being merely a clever artifice that requires correction in order to account for the peripheral situation (Hudson, 1997, p. 7@). &he only significant sense in which it can be useful is in the insight it provides with respect to the identity of the likely winners in interstate competition. A 'likely winner' in the periphery is a state that minds its own business, concentrates its attention on trade and development, and abides by the rules of the game set by the great powers. But, then, this means that a key structural feature that 'drops out' of Walt's formulation needs to be recovered: the attributes of states (pp. 7@9). And, if we look closely, we come to the realisation that yet another feature, namely, the 'organising principle' theorised by Walt, also requires

alteration in the case of peripheral states, to accommodate an incipient and imperfect hierarchy (p. 20). This comes as no surprise, since Walt's original approach aimed solely at an understanding of great power politics (Resende Santos, 1997, pp. 155). Following Morgenthau, the preference of 'peripheral realism' is for a normative theory. Unlike the classical realist, the focus of policy lies on the citizens, not the state itself. Following Walt, understanding 'structure' can be profitable, but unlike Walt a peripheral realist also needs to look at the unit level.

That, then, are the substantive arguments and prescriptions of peripheral realism. The core of the programme derives from a reformulation of structural constraints and mechanisms that reproduce them.

The structure of the interstate system is better characterized by the concept of an incipient and imperfect hierarchy than by the anarchy postulated by realist theorists. This incipient hierarchy, in turn, is enforced by sanctions in places that are not always effective but that make it costly for weaker states to challenge the strong, especially when evaluating costs from a citizen-centric perspective (Schudson, 1997, p. 160).

Studying structure from the perspective of unit-level features implies the acknowledgement of a certain functional differentiation between great powers and weaker states, but this is said to be a matter of degree along the continuum that empirically exists. Because of the differentiation, economic issues become a crucial tool of dominant states, which are in a position to link crucial economic issues to desirable political attitudes on the part of the weaker states (Schudson, 1997, p. 159). The functional differentiation requires us to divide states into three types: great powers (states that command) and weaker states (states that obey and rebel states). The hierarchical arrangements are not necessarily dyadic, unlike Easley's formulation. One can identify this asymmetry across the whole system. Within each different category, of course, the relations are more symmetrical, all things being equal. Differing from Walt, Schudson (p. 160) ascribes a potentially destabilizing role to 'rebel states'. They are the equivalent of outlaws or mafias in domestic societies. Rebellious states are punished by this hierarchical and functionally differentiated structure. Citizens suffer much more from these punishments than the ruling elites. Obedient states, on the other hand, often reap short- and long-term benefits of cooperation and compliance which translate, as far as possible, into peace and well-being of the citizens. Great powers can afford the luxury of adventurous military exploits, but weaker states should treat development as priority and be more aware of the mercantilistic link between power and wealth.

(pp. 109-110).

Being obedient pays off in the periphery. Realising down the 'black box' of the state as 'unitary' actor, the author advises against four basic pursuits. First, a peripheral state must abstain from interstate power politics and devote itself to local economic development. Secondly, it should avoid costly idealism and only go after good causes abroad if there are no risks and costs involved. Thirdly, this type of state should abstain from risky confrontations with great powers when they engage in policies that are detrimental to universal good causes but that do not affect the peripheral government's material interests. Finally, a peripheral state willing to obey the great powers and reap the benefits of peripheral realism should abstain from unproductive political confrontations with them (Hudson, 1997, pp. 109-110). These principles, I say, truly satisfy the metatheoretical requirements stipulated, after Morgenthau, of the application of a consequentialist ethics to the formulation of foreign policy (see p. 15).^{9G} Moreover, they derive from the metatheoretical avoidance of the 'anthropomorphic fallacy', since a line is clearly drawn between these policies and what would otherwise be a power politics pattern of interstate behaviour (see also Hudson, 1997). Finally, they are not only consistent with, but also enabled by, Hudson's critical reformulation of Walt's conception of 'structure', the third metatheoretical principle at stake (1997, pp. 111-112). For all its consistency, though, peripheral realism ultimately fails to designate concrete criteria for testing the theory. One could infer, from the notions employed by Hudson to judge the perverse effects of mainstream IR approaches in Argentina and other weaker states, that the social consequences of policy derived from peripheral realism would be part of the test. Very well. Hudson's peripheral realism informed, in fact, the foreign policy of so-called 'carnal relations' with the * during the administration of President Carlos Menem (Antoro, 1999, pp. A5). The central normative point of peripheral realism has to do with citizen welfare and economic development. Menem's domestic and foreign policies are commonly blamed for the economic and political disaster that ensued few years later (Gervó, 1999, pp. 112-113). Hudson's (1999, esp. conclusion) reply is that he never endorsed economic 'neoliberalism', but rather a 'mercantilist' policy. And he writes as an *ad hoc* additional hypothesis to the theory already in circulation before the Menem regime. Whether Menem's economic liberalisations were indeed the cause of the disaster would demand,

^{9G} If there is any question of that, see Collyer's (1999, pp. 9-10) discussion of Morgenthau's *dictum* that 'political ethics judges action by its political consequences' (cited and discussed on p. 9A).

of course, an economic study on its own. #scudJ's reply, however, is by no means straightforward. &here are those who, li'e -ans'-ermann -oppe, would blame (enem for *too much* statism, *not* liberalism. &he attempt to 9ustify -oppe's peculiar anti-mercantilist view of the wealth'power ne\$us is our ne\$t case.

Deducti#e theory! time3preferences! -ealth and po-er

A series of highly controversial essays in -ans'-ermann -oppe's *<emocracy 5 The Hod that Failed* pioneer the application of 'Austrian' political economy to the theoretical comparison between democracy, monarchy and mar!et'anarchism. In passing, they briefly deal with cycles of international hegemony and decline from an economic perspective. As far as I understand, this is the first time -oppe's views on world politics are compared to !ey IR te\$ts on hierarchy in a metatheoretical study. -oppe's normative inclinations, consistent with the ethical perspective advanced by leading Austrian school economist (urray Rothbard (.99@), favour anarchism or 'natural order' over 'monarchy', and the latter over 'democracy'. &he !ey value is to rule out any initiation of aggression, the state being by definition a ma9or initiator of aggression. -oppe's (.999) defence of mar!et'anarchism includes the claim that even protection services would be sold, rather than financed via ta\$ation (an instance of aggression). In -oppe's view, surprisingly, monarchy is inherently less e\$plotative and violent than democracy.^{9A} 3roving the point is not only a way of challenging the democratic hegemonic pro9ect that legitimates the *status \$uo*, but also to e\$plain the desirability of less e\$plotation and aggression beyond the ethical aspect (-oppe, .99/).⁹⁼ A political system organised along these lines would lead to more productivity, prosperity and even higher cultural or 'civilisational' standards. In -oppe's conception, there is a positive relation between 'degree of civilisation' and the diffusion of the 'non'aggression principle' in any given society. A line, then, is clearly drawn between e\$plotative:uncivilised systems and their free:civilised counterparts (see -oppe, 5//., p.> where the distinction is employed at the personal level). &here are scattered references to international politics and cycles of hegemony and imperialism in the international system. &his is where we find an implicit theory of hierarchy.

9A -oppe means ideal types of 'democracy' as publicly'owned government and 'monarchy' as privately' owned government. In his definition, the logic and structure of post'.9th century monarchies is closer to the 'democracy' type, especially after .9.G.

9= 6hallenging what is perceived as the ideology legitimating the status "uo (institutionalised initiation of aggression via state power) is a !ey way to pursue change, since -oppe re9ects violent populist revolutionism (see -oppe, 5//., pp.9/G).

The Austrian school research programme that has been more inclined to extend the school's framework beyond economics is *praxeological* (Rothbard, 1962, p. 1).^{9>} In this tradition, praxeology is a metatheoretical approach that favours analytical and deductive thinking and reflects the mainstream empiricist conception of social theory (see Colander, 1997).^{9@} Its cornerstone is the fact of human action, regarded as the fundamental starting point for every social theory (and not only economics). Human action in this technical sense is purposeful behavior (Mises, 1999, p. 1) or, in other words, an actor's purposeful pursuit of valued ends with scarce means (Mises, 1962, p. 17). This is regarded as a solid starting point by virtue of being true *a priori* if we want to prove there is no such a thing as human action, we will be engaging in purposeful behaviour, thus making the attempt self-referentially inconsistent (Callahan, 1976, p. 59). Praxeology as a metatheory involves assuming the 'action axiom' and building up an edifice of theorems deduced from this starting point (Mises, 1962, pp. 1-2). This is the metatheoretical 'drive' behind Mises's efforts. Contrary to mainstream theorising, conclusions obtained this way can be *illustrated* by historical data, but historical data can neither *establish* nor *refute* them (Mises, 1962, p. 18). In fact, says Mises (p. 16), 'someone who wanted to "test" these propositions should be regarded as confused, for deductive social theory trumps and corrects experience, and not the other way around. Praxeology, therefore, is everything a good positivist claims one cannot and shall not be interdisciplinary, theoretically oriented, and dealing with both positive empirical *and* normative questions' (p. 14). Now exactly does this apply to the notion of hierarchy in the international system? Before we get to that point, it is necessary to follow the 'deductive chain'.

The first step in the argument is the definition of 'time preference' as a phenomenon permeating all human action. Since the point of acting is to pursue 'valued ends' given scarcity of 'means', and since 'time' affects both the action itself and the duration of serviceability of those means, then we may postulate that present or earlier goods are, and must invariably be, valued more highly than future or later ones. The implication, delaying one's consumption of a present or earlier good is only likely to

9> Alternative research programmes in the Austrian tradition are the 'spontaneous order' approach of, e.g., Hayek and the 'hermeneutical' approach of Eudwig Eachmann.

9@ & here are other approaches that claim to be praxeological, but in different senses. The philosophy of Zygmunt Bauman, the strategic thought of Claude Lévi-Strauss and the international theory of Raymond Aron employ the term, but by far (judging by volume of publications) the most known consistent use as a school of thought is that of the 'Austrian economics' of Eudwig von Mises and Murray Rothbard.

happen if there is a perceived incentive to do so, such as for example a future 'premium' for present savings. Low time-preference rates contribute to capital accumulation, increase in productivity, and enjoyment of more 'scarce means' at the end of the process of production. High time-preference rates do the opposite. Like a child, one tends to opt for instant or minimally delayed gratification (Oppenheimer, 1997, pp. 1-2). Institutional factors contributing to lowering the overall time-preference rate include a stable monetary system (no inflation), property rights and so on. The process of civilisation emerging from this overall decrease in time-preference rates is proportional to the lowering of 'interference' with voluntary exchange and appropriation from nature. Any kind of recurrent violence is but mainly government institutionalised initiation of aggression via taxation leads people to a more 'present-oriented' pattern of behaviour and represents a tendency toward decivilisation (pp. 1-2).

The second step of the argument stipulates that not all types of government systems affect time-preference rates equally. There is a qualitative distinction between the way 'monarchy' (privately owned government) and 'democracy' (publicly owned government) affect these rates, on at least two accounts:

- (4) A private government owner will tend to have a systematically longer planning horizon, i.e., his degree of time preference will be lower, and accordingly, his degree of economic exploitation will tend to be less than that of a government caretaker; and (5), subject to a higher degree of exploitation the nongovernmental public will also be comparatively more present-oriented under a system of publicly owned government than under a regime of private government ownership (Oppenheimer, 1997, p. 6).

Because of the way this affects a ruler's longer-term expectations, there is, structurally speaking, much more incentive for a low time-preference rate when a country's government is 'privately owned' (as in pre-19th century Europe) than 'publicly owned' (as in contemporary democracies in the West). Oppenheimer employs a good number of key secondary sources in international historical sociology to illustrate the point, looking at 'indicators' of governmental exploitation (e.g. taxes, government employment, inflation rates, public debt, legislation/regulation) and 'indicators' of present-orientedness (e.g. high interest rates, low contributions to charity, crime rates). Such 'domestic' indicators illustrate the revisionist conclusion that the historic transition from monarchy to democracy represents not progress but civilizational decline (pp. 1-2).⁹⁹ I say 'illustrate', because Oppenheimer himself has already asserted that the theory can only be

⁹⁹ The author makes reference, for example, to the works of Martin van Greveld, Gerard Laidel and Charles Kilby.

'tested', or 'verified:refuted' in terms of its internal consistency. While not immediately linked to my position, the general principle established in this step of the argument is essential to the conclusion. The principle consists in connecting economic liberty to prosperity and 'civilisation' via time-preference rates.

The next step of the argument, or the '13#' portion, comprises a number of scattered passages which I now connect in order to 'distil' the theory of hegemony cycles embedded in them. From the violent nature of the state as a 'legitimised monopolist of expropriation' in a given territory, we can infer its tendency to domestically expand and centralise (given lack of serious internal competition). We can, moreover, infer that different states will compete for more control over territory and populations, engaging in eternal expansion as well (pp.55F).¹¹ The degree of 'economic liberty' determines not only the 'degree of civilisation' in a certain country, but also by implication the amount of resources a government has available to pursue further centralisation and expansion. The logic of expansion often leads to a paradoxical cycle of hegemonies. This is also a step-by-step 'praxeological' argument.

The international aspect of the expansionist drive of states involves, of course, competitive elimination. The move is from systemic fragmentation, in which it is relatively easy for an exploited population to 'vote with their feet' (thus constraining domestic incentives for government to exploit even more), to systemic concentration (-oppe, 5//., p.127). As Alexander Wendt (5//F) once observed (but for different reasons), there is a long-term tendency to centralisation and decrease in the number of states in the system. Ultimately crucial for whether a state will expand or eventually disappear is the relative amount of economic resources available to it (-oppe, 5//., p.127). But, then, the 'freer' the state, the more 'prosperous' it is%

¹¹ This expansionist drive and, more importantly, the means whereby it occurs, also depends on whether a country is 'privately owned' or not. Democracies tend to be much more destructive when they go to war, while the constraints on adventurous dynasties are legion. Here, again, the distinction is not a matter of degree, since the 'civilizing' forces in a monarchical state are insufficiently strong to overcome the fundamental, countervailing tendency toward falling time-preference rates and ever-expanding ranges of private provisions. In publicly-owned governments, conversely, the 'civilizing' effects of government can be expected to grow strong enough to actually halt the civilizing process, or even to alter its direction and bring about an opposite tendency toward decivilization. In -oppe's somewhat prejudiced metaphors, La progressive infantilization and brutalization of social life via capital consumption and short-sighted horizons and provisions (-oppe, 5//., pp.117G). Besides the decision to go to war, the mode of warfare is also influenced by 'regime type'. Thus, in the transition from 'monarchy' to 'democracy' we can observe a change from limited warfare to total war (-oppe, 5//., pp.129; see 5A=). In light of what we may deduce from the incentive structures in each type of government, monarchies are overall less violent, exploitative and more civilised and self-restrained than democracies. This is as close as we can get to a (anarchic) peace theory.

As other things being equal, the lower the tax and regulation burden imposed by a government on its domestic economy (O), the larger the amount of domestically produced wealth on which it can draw in its conflicts with neighbouring competitors (...). States which tax and regulate their domestic economies little or liberal states tend to defeat and expand their territories at the expense of non-liberal ones (Hoppe, 1993, p. 5).

The first paradox in this cycle of hegemony, therefore, is that states on top of the hierarchy, which are the most internationally aggressive states (such as colonialist powers), are domestically the most 'liberal' ones in the system. This, however, is also their Achilles' heel, and another paradox.

As the further the process of more liberal governments defeating less liberal ones proceeds or i.e., the larger the territories, the fewer and more distant the remaining competitors, and thus the more costly international migration or the lower a government's incentive to continue in its domestic liberalism will be. As one approaches the limit of a one world state, all possibilities of voting with one's feet against a government disappear (p. 5; see also 5F.).

Carl Schmitt's (1923, p. 16) argument that 'liberal' states need a great protection of state power to 'make markets work' is an ironic statement well known in IS#. Hoppe's praxeological theory of international hierarchy puts Schmitt's logic upside down. (After liberalism in fact works so well that governments decide to use its resulting prosperity to expand and eliminate external competition. When the state grows, though, the incentive structure that would keep it restrained and 'liberal' also changes, so in the long run it tends to become centralised, exploitative, driving time preference rates up and productivity down. As a result of 'killing the goose that lays golden eggs', the hegemon eventually declines.

Hoppe's constant reminder that the theory is purely deductive and therefore not open to 'empirical test' often jeopardises the persuasiveness of the argument to those too attached to mainstream theorising. The fact that he omits links in the 'deductive chain', considered unproblematic (having been previously established by other Austrian school economists like Ludwig von Mises and Murray Rothbard) also weakens the argument from a rhetorical perspective, particularly in the case of those who are unaware, or dismissive of, the approach. Despite this, Hoppe still opts for a way of theorising that is consistent, as far as possible, with the principles of praxeological metatheory. The 'test' for the theory is to find logical fallacies in the argument. If we assume this point of view, then a helpful exercise for an 'Austrian' scholar would be to include an appendix

after each study containing a formalisation of the argument in predicate calculus. &he use of analytical philosophy would be an essential tool for the 'Austrian' economist, yet to date no well-known attempt has been made to formalise economic praxeology in this sense, let alone -oppe's view of the hegemonic cycle. Besides, even in the absence of further specification for theory test, we can still ask, for example, whether economic liberty is indeed the main cause of international hegemony, for an usually deductive claim could be made that would incorporate 'technology' or some other unrelated possible cause. &his sort of critique is more reasonable than some of the 'eternal' claims often made that praxeological reasoning is pseudo-scientific because it does not bow to empiricist standards stipulated by the mainstream (e.g. Lunge, 1999, pp.55A'). Eile #scudJ, -oppe denounces the exploitative interests of the ruling elites and draws a line between governments and citizens. Unlike #scudJ's account, -oppe's view of the relation between wealth and power challenges by implication the mercantilist assumptions of realism.

Eurocentrism and hierarchy in international thought

In his study of 'Eurocentric' assumptions in international theory, -obson interprets a wide selection of texts over a long time-span (1850-1950). &he main critical conclusion is that what we encounter in the vast majority of international theory is the provincial or parochial normative purpose of *defending and celebrating the ideal of the "best in world politics"* (-obson, 1955, p.55, original emphasis). &his is enabled by Eurocentric narratives operating in different ways depending on the text. In any case, their way of shaping international theory (and particularly the IR discipline) leads it not so much to explain international politics in an objective, positivist and universalist manner but, rather, to parochially celebrate and defend or promote the West as the proactive subject of, and as the highest or ideal normative referent in, world politics. IR theory, therefore, functions as a vehicle, or repository of Eurocentric metanarratives (pp.15). -obson's study synthesises several manners in which this occurs, always looking for evidence in key historical and contemporary works. 'Eurocentrism' has four basic variants, depending on whether it manifests itself as an 'institutionalist' argument or a 'scientific racist' claim, combined with a 'pro-imperialist' or an 'anti-imperialist' stance. Paternalist Eurocentrism is an imperialist form of institutionalist claim. Offensive Eurocentrism is, at the same time, racist and imperialist. Anti-paternalist Eurocentrism is institutionalist and against imperialism. Defensive

#urocentrism combines an anti-imperialist stance with scientific racism (pp.F'@). &here are also variations in the presentation of #urocentrism. Before 1990 international theory was overtly #urocentric, with the West imbued with purely virtuous and/or progressive properties that in turn lead it to pioneer all that is progressive in world politics. During the Cold War, it employed sanitized terms like 'modern:traditional', 'core:periphery' rather than 'civilisation:barbarism' 0 a covert or subliminal type of presentation. A more overt approach has been recovered since the end of the Cold War, although critical IR theory, in the backseat of the discipline, kept traces of #urocentrism in the covert format (pp.9'./). Based on this classification, we should expect the theories analysed in this chapter to fit into the post-1990 categories.

Crucial for our understanding of the #urocentric character of IR theory is the fact that, at least implicitly, it is embedded in assumptions about hierarchical relations between different groups of states. International theory conceptualises sovereignty in the anarchical system only as a derivation of an underlying worldview which has hierarchy in the centre, based on a presupposed standard of civilisation. For all its disguise under the rhetoric of 'sovereign equality under anarchy', IR theory is predicated on the unequal field of global:civilizational hierarchy and graduated sovereignties. &his can be translated in practice as either formal or informal hierarchy. Formal requirements make non-Western sovereignty conditional upon meeting the 'civilisational standards', whereas in the informal case its self-determination is maintained, but with an incentive to assimilate or culturally convert to Western civilizational norms (Robson, 2005, p.9). In both cases, reflection on unequal sovereignty results from a transposition, into the international level, of the 'line' drawn between the civilised and the uncivilised via domestic analysis of institutions, culture and/or race. Regardless of the theoretical approach, those units deemed to be domestically 'civilised' are accepted in IR scholarship as more capable of international political agency than those which are not (pp.FFGff).

Before we apply Robson's metatheoretical analysis, a critical comment on its main shortcoming is in order. He purports to uncover the #urocentric normative bias of international theory, which is often implicit, and to draw a darker picture of the discipline of IR. &his is paramount, first, because otherwise we will continue (O) to reproduce this discourse of power through our own writings, but also because international theory is inherently politically performative (Robson, 2005, p.=). However, Robson admittedly selects for analysis texts which will corroborate and

illustrate his claims, and acknowledges there are contemporary non-Western exceptions, which he prefers to leave aside (p. 11). Moreover, he also does not consider non-Western texts written before the institutionalisation of an IR discipline, perhaps in non-European contexts. If international theory is politically performative, and if seeing our discipline without discerning its Western bias further reproduces it, then surely *metatheory* is also politically performative, and the portrait of IR as a large set of 'Western discourses' would further marginalise non-Western perspectives. If there ever was any chance of highlighting non-Western discourses as part of the conversation, Johnson had it, but decided to merely assert, in the conclusion, the need to work out how a non-Western foundation for IR theory might be reconstructed (p. 12). While I expect this type of research to be the next step in an already long list of metatheoretical studies on 'non-Western IR', I am unaware of any other study extending Johnson's approach to alternative material such as 'peripheral realism'.

We may start with Johnson's own analysis of neorealism, a research programme denounced for its residual paternalism due to Walt's conjecture that 'hegemony has a 'benevolent' character' (Johnson, 1995, p. 10). Here Johnson may have a point, especially if Walt's conjecture is read against the context of the post-Cold War literature on the so-called 'unipolar moment' (e.g. Rothhammer, 1997). However, the criticism applies to an *ad hoc* auxiliary hypothesis, not to the bulk of the theory. Issue may also be taken against Johnson's (1995, p. 12) interpretation of systemic change and functional similarity in *Theory of International Politics*. It is said that Walt denies that structural international change is possible, when in fact he allows for large-scale change, but narrows it down to an alteration in the distribution of capabilities across the system, making the 'ordering principle' and the 'functional differentiation' of units drop out of the equation.¹ Furthermore, in the commentator's understanding, neorealism is supposed to apply to *all* states, regardless of their cultural foundations, or identity, whether they be Western or non-Western (original emphasis). This is only half true, because cultural factors are indeed denied salience in the theory, but Walt makes it clear that neorealism is an explanation of the way relevant political units operate internationally. Therefore, the 'universality' of the theory applies only to *relevant* states in similar conditions.

While these claims are unpersuasive in Johnson's reading of neorealism, the

¹ Obviously, Walt's 'endogenous' account of change may be, and has been, criticised from many different angles. But this only reiterates the point that *there is* such account in the first place (Ashley, 1995; E. 2. Jönsson, 1995).

remaining critiques are more pertinent. -e correctly points out that Walt's 'theoretical fiat' isolating the international realm discards the possibility that hierarchy of any sort can exist in the international system. As a result, this simplification, sharply distinguishing between anarchy and hierarchy, occludes the possibility of revealing various hierarchical international political formations, thereby 'sanitising' colonialism and other deleterious practices from the account. Eurocentrism, here, would be an interpretation of such asymmetrical relations via a Western liberal framework that assumes the formal sovereign equality of states. Another point of Eurocentrism is Walt's denial of Eastern agency, such as for example the unwillingness to admit that even with the disproportion between the US and the rest in the Cold War, Vietnam and Afghanistan still thwarted the plans of the superpowers (Johnson, 1975, pp. 5/6-9). Johnson's application of his framework to neorealism in order to detect traces of 'Eurocentrism' is only a partial success. (Maybe his rush to critique before establishing a close reading of the text is a defect inherent to general metatheoretical/historical studies covering large samples and a longer time-span. A suggested refinement to the approach would be to 'dig deeper' into individual texts and pay more attention to details. For example, it has been suggested that another source of Eurocentric bias is the fact that Walt draws his concept of structure from the classical anthropology of Bachelard and Evans-Pritchard, uncritically transposing the original delineation of 'civilised' and 'traditional' societies into the international realm (Johnson, 1975).

It is surprising that Johnson's study did not consider Eyal's 'dyadic' theory as one of the samples. The model is one of the few to place hierarchy at the centre of IR theory, besides being increasingly popular in the recent IR and IJ literature. Eyal's formulation attempts to provide a 'neutral' cost-benefit analysis of dyadic hierarchical arrangements from both points of view of dominant and subordinate states. Nevertheless, it often slips out normative statements like 'With great power comes great responsibility'; 'the right to punish noncompliance ultimately rests upon the collective acceptance or legitimacy of the ruler's right to rule'; or 'the key problem in any hierarchy is limiting abuses of authority by the ruler' (Eyal, 1979, pp. 10, 11, 12). From Johnson's critical perspective, this sort of statement typically assumes a Western liberal political narrative, but naturalises it under the guise of 'scientific neutrality', on the one hand, and 'consent of the ruled' on the other. In Johnson's story, accounts of unequal sovereignty are to be expected in theories with this kind of bias. Eyal's theory not only postulates unequal sovereignty, but also intrinsically connects it to international

hierarchy. All sovereignty is divided, more or less. Hierarchy is simply the counterpart to this variable sovereignty (Ea!e, 5//9, p.A.). Hobson's identification of a new form of Eurocentric paternalism which, after the Cold War, adopts an overt defence of Western (especially *) interventionism in world politics, can be matched to Ea!e's central claim. says Hobson (5/.5, p.5A>)%

The Western liberal wing of mainstream international theory relies on a paternalist Eurocentrism that sings the world into existence with the idiom that 'things can only get better', and that through paternalist interventionism the East can be culturally converted along Western civilizational lines in order to make the world a better place for all peoples.

And here is Ea!e's portrayal of the ideal hierarchical arrangement that would effectively work according to his theoretical predictions%

Both dominant and subordinate states have to be better off in hierarchy than in strictly anarchic relations for the contract to be fulfilled. Subordinates must yield some portion of their sovereignty and accept their status as legitimate, appropriate, and perhaps even necessary. In return, they receive a political order that allows them to escape in part the state of nature that is international anarchy. Dominant states, in turn, get to set the rules of their international orders in ways that benefit not only others, but also themselves. Yet, at the same time, they must bear the governance costs of producing those orders, disciplining subordinates, and credibly committing not to misuse their (*sic*) authority they earn (5//9, p.9F).

Notice Hobson's elements here% paternalism that 'sings the hierarchical world order into existence', with the effect of a considerable improvement. The 'cultural conversion' is, however, only implicit, in the idea of a Western liberal 'contract' including the acceptance of hegemonic 'legitimacy' escaping from the 'state of nature' of anarchy (very similar to the liberal narrative for the domestic social contract). Besides, indicators of 'Eurocentrism' are also present (i.e. passive subordinates, active hegemons)% subordinates must yield and accept hegemons, and receive order. Dominant states set the rules and make order. They are also active when they heroically bear the governance costs and commit to use well the authority that they learn. Standing Hobson's account to Ea!e's dyadic theory further corroborates the narrative. This, however, is to be expected, given Ea!e's theoretical and academic influences.

Peripheral realism, on the other hand, seems to be a more difficult case. This other perspective assumes hierarchy, even in the formal sense, as a given, and disputes any attempts to 'unrealistically challenge' this structure. Moreover, the policy implications of this approach are very similar to those defended by Ea!e, and both could

be politically criticised for being e\$aggeratedly supportive of the status "uo. -owever, peripheral realism is written from the perspective of statecraft in weaker states, even 'non'7estern' states in -obson's sense. #scudJ denounces Lthe ethnocentric "uality of the international relations theory coined in the *nited tates and elsewhere in the Anglo'American worldM, but specifically targets those in Lmisguided &hird 7orld governmentsM who LmisreadM this type of theory Lwithout a critical e\$amination of whether the assumptions of these theories are adaptable to local circumstancesM (.99>, pp.A'=). &hat is to say, it is *precisely* its 'non'7estern' character that ma!es room for peripheral realism to modify classical and structural realist approaches of (orgenthau and 7alt? and to ma!e them wor! for weaker states. It is *precisely* the assumption of hierarchy and ine"uality, challenging (not corroborating) mainstream forms of realism, that enables it to theorise foreign policy deemed to be better suited for a 'non'7estern' situation (see &ic!ner, 5//Fa, 5//Fb, 5//@). ,or -obson, an easy way out would be to blame the influence of '7estern' IR theory on #scudJ's thin!ing, but this would not do 9ustice to his deliberate re9ection of it *on account of* its '7esterncentric' features. Another solution would be to label peripheral realism '#urocentric' because of its '7estern'liberal' assumptions on the centrality of citi?en welfare. -owever, this argument alone would not do. ,or it is a central claim of the theory that '7estern IR', being state'centric, merely pays lip service to these 7estern'liberal principles, while peripheral realism is truly politically liberal. A better way to deal with #scudJ's case from -obson's perspective would be to analyse peripheral realism's tripartite hierarchical arrangement between 'states that command', 'states that obey' and 'rebel states'. -ere we find a reproduction of claims about 'standards of civilisation' which are in -obson's conception at the core of a typical '#urocentric' approach (see also 7all!er, 5//=). #scudJ has to challenge 7alt?'s 'parsimonious' assumption that weaker states are irrelevant at the systemic level. In contrast, it is argued that rebel states do indeed disturb the system. &he claim is not anymore that 'rebel states' are illiberal and their governments 9eopardise their citi?ens. Instead, these less 'civilised' states are labelled as 'troublema!ers' and are 'tamed' by the 2reat 3owers via sanctions and interventions in order to stabilise world politics. #Stended to the odd case of peripheral realism, therefore, -obson's narrative seems to apply, although the analysis holds only after an e\$amination of #scudJ's theory in its own terms.

&he application to -oppe's pra\$eological political economy is more straightforward, since the author himself freely employs '#urocentric terminology and

the distinction between 'civilisation' and 'barbarism'. The twist here, though, is that -oppe as an individualist anarchist denies any legitimacy to the state. (Moreover, the theory postulates an inverse relation between the degree of 'statism' and the degree of 'civilisation'. For this reason, any institutional eurocentrism (according to -obson's conceptualisation) must be of the non-statist sort very different from most theories interpreted in -obson's study. In -oppe's work there is, indeed, a deliberate attempt to vindicate libertarian individualist market anarchism as a more 'civilised' way of life. This is intrinsically linked to the relation between relatively lower time preferences within such institutional framework, on the one hand, and indicators of 'prosperity' on the other. The standards of economic and cultural behaviour associated to the 'ethics of capitalism' described as 'civilised' are emphatically based on a eurocentric narrative about the historical formation of a 'prosperous West' enabled by such virtues. The author, however, denounces the growing centralisation and eternal expansion of 'prosperous' states enabled by a 'liberal' wealth/power nexus. In fact, ideally local communities and even individuals should consider delegitimising state coercion and support secessionist causes everywhere (-oppe, 1995, pp. 19; 5=95). The policy, of course, is a denial of the Western hegemonic discourse of 'Liberal' democracy. With this caveat, we may consider this theory an institutionalist, but anti-imperialist, eurocentric approach in -obson's scheme.

Final remarks: levels of metatheoretical analysis

(Meta)theory operates in this chapter in a number of different ways, and this series of final remarks purports to clarify them. We moved from the historical debate on the role of Westphalia as an IR narrative of 'hierarchy' to 'anarchy' transition to a theoretical debate of what 'hierarchy' means. In the field, the concept of hierarchy denotes some sort of asymmetry in the international system. It is actualised in distinct manners by each of the theoretical approaches analysed here. These theories, in turn, are 'driven' in different ways by metatheoretical principles. In terms of a metatheoretical framework designed to understand hierarchy, -obson's critique of eurocentric assumptions in IR and IR theory provides a set of tools. Drawing on this formulation, it is possible to interpret some of the normative elements behind each of the theories selected for our analysis.

Realism is, at the same time, 'instrumentalist' and 'structuralist'. Structure is postulated as a way of creatively distinguishing the international realm of investigation

from domestic political systems. Hierarchy characterises domestic politics and anarchy describes international political interaction. Most explanations and predictions in the theory follow from this combination of instrumentalist simplification and systemic theorising. Other claims are made with reference to post-Cold War hegemony and balance-of-power. These claims do not follow the same metatheoretical standards stipulated by neorealism as a research programme. The mainstream empiricist approach to the study of dyadic hierarchy is presented as critical realism and in tension with 'positivism'. Upon scrutiny, its empiricist character has been uncovered, and we may correctly locate it closer to 'positivism' than to critical realism. Hierarchy means a relation of authority and consent between, respectively, a dominant state and a subordinate state. Both neorealism and the dyadic approach are state-centric, but the former makes a point of 'systemic' theorising, while the latter incorporates a number of 'reductionist' (i.e. unit-level) variables. Both are Eurocentric and 'paternalist' in Hobson's understanding, but Hobson's interpretation of neorealism may be questioned in a more detailed reading.

The other two theories of hierarchy are less focused on the state or, rather, negatively focused on the state. 'Peripheral realism' metatheoretically criticises the normative and logical problems in the 'anthropomorphic fallacy' of the state-as-unitary actor model assumed by mainstream IR theorising. It places hierarchy at the core of the framework in order to provide advice on foreign policy for a 'weaker state'. 'Rebellious states' are denounced due to the fact that their risky confrontations of the Great Powers lead to sanctions and, consequently, harm to the population they are supposed to protect. 'States that obey' are portrayed as the model of successful and prudent foreign policy in the periphery. Peripheral realism fails to devise clear evaluation criteria of its own, so we have to derive a political critique from its dealings with mainstream theories. Since the main criterion has to do with the consequences of using such theories, if we find questionable outcomes related to the use of peripheral realism, then there is reason to mistrust it.

Another view of hierarchy reflecting state-centrism is the theory of cycles of hegemony and decline implicit in 'praxeological' deductivist political economy. In this view, hierarchy begins at the individual level, and inter-state asymmetry is only a derivation from it. The theory connects 'civilised' low time-preference rates of individuals to 'prosperity'. The aggregation of these rates are encouraged by a high degree of economic freedom in a country. A 'liberal' country manages to amass

sufficient wealth to become more centralised, exploitative, and imperialistic. It makes sense to assume an international asymmetry in power and command, especially of the colonialist sort. However, these expansionist and centralising moves, in the end, increase the aggregate time-preference rates and the same country declines. This view of the hegemonic cycle is the culmination of a 'deductive chain'. The philosophical rejection of 'empirical testing', and affirmation of an 'internal test of coherence', leads us to ask why no effort is made to formalise the claims in order to facilitate logical testing. If peripheral realism, this approach requires a refinement in the metatheoretical criteria for theory evaluation. But if we employ Hobson's framework, we can see that peripheral realism is 'covert' in its Eurocentric character, while philosophical political economy, at least in this instance, is 'overt' in its adoption of categories like civilisation:barbarism.

This chapter highlights the operation of metatheory as a constraining and enabling conceptual 'mechanism' in a number of ways. The first level is 'immanent' to the approaches selected for analysis. Because each of the main authors state their own metatheoretical views, the initial step is to clarify their influence on the theorisation of hierarchy. Here we see different 'drives' behind each approach: instrumentalist and systemic (neorealism), empiricist (dyadic hierarchy), normative (peripheral realism) and deductivist (philosophy). This accounts to a great extent for the particularities in each of the 'ways of theorising' adopted. An additional step is to evaluate the internal consistency of the relation between metatheoretical preferences and the theoretical claims in each case. Neorealism and dyadic hierarchy are not fully consistent with their metatheoretical claims. Peripheral realism and philosophy require further specification on theory evaluation, but would also, at first sight, appear to fall short of their own standards. The first level, therefore, is 'immanent' to each approach. It is, moreover, an Internal: Intellectual exercise in which metatheory as clarification (hermeneutical) and metatheory as mechanism of evaluation (evaluative) and/or refinement (corrective) are at stake. The second level cuts across the approaches. It consists in the use, by each of the authors, of their own metatheoretical standards against opposing views of hierarchy in the field. Speaking from a 'systemic' perspective, neorealism dismisses 'reductionist' theory as not parsimonious enough. Mainstream empiricism in the dyadic approach refuses to take into account 'ideational' or 'normative' views of hierarchy because they are not operationalised or measurable. Peripheral realism, in its ethnocentric and 'non-Western' normative perspective, provides a critique of the 'anthropomorphic fallacy' in

mainstream IR theorising. Here, not only Internal:Intellectual elements are at stake, but also Internal:Contestual (theory:evidence, theory:politics) and External:Intellectual (philosophy:theory). (Meta)theorising as critique and as theoretical evaluation are the main occurrences.

The third and the fourth levels have to do with the introduction of Hobson's overarching framework employed to spot 'Eurocentric' narratives and normative commitments in international thought. On the one hand, this metatheoretical framework operates as an External:Intellectual (philosophy:theory, theory outside IR:theory in IR) and Internal:Contestual (theory'in'society) mechanism of clarification, disciplinary history and critique. These three ways of meta-theorising are highlighted and declared by the author himself. The study is, on the other hand, an analysis motivated by the harmful impact of 'Eurocentric' theory on political practice (critical), an account of the formation of IR that stresses this aspect, often absent from traditional accounts like the 'Great Debates' narrative (historical) and, finally, an interpretation (hermeneutical) of more than two centuries of international thought with special reference to IR and IR theory (Hobson, 2002, pp. 1-10). The interpretation uncovers the inner Eurocentric normative narratives 'driving' theorisation. Most of IR theory is said to be based on implicit views of hierarchy rather than anarchy. This is the third level. The fourth and final level at which metatheory operates in this chapter is that of my own evaluation of Hobson's framework. In this sense, it is a metatheoretical analysis of Hobson's metatheoretical approach. It is a meta-metatheoretical discussion, mostly of the Internal:Intellectual variety. I test Hobson's claims against neorealism and find some of them wanting (evaluative meta-theorising). I also point out how they could be improved (corrective meta-theorising). An indication is left that the author's views should be extended to 'non-Western' theories, lest it reify a 'Eurocentric' IR discipline (critical meta-theorising of the Internal:Contestual sort). In extending Hobson's analysis to peripheral realism, I claim it still applies to this difficult case of 'non-Western Eurocentrism' (corrective meta-theorising), and so it does to post-colonial political economy. It is this meta-metatheoretical analysis that concludes the illustrations of the general claims advanced in this thesis. We move now to a summary of my contributions and the general conclusions.

Conclusion

The present thesis may now be concluded. After the problematisation and conceptualisation undertaken in the first part, the thorough examination of IR views on the matter in part two and the 'illustration cases' in the final part, two types of final remarks are in order. The first addresses the original contributions of this study, both primary and secondary, exploring avenues for further research. The second type of final remarks highlights the answers provided to the issue at hand and suggests criteria to test my claims.

A summary of the primary contribution

This study has investigated the role of metatheoretical scholarship in the academic discipline of IR. Unlike the literature on the topic in our discipline, which treats the issue only in passing, its main original contribution is the treatment of the topic in general terms, formulating a qualified defence of metatheorising. First, the thesis provides an extensive conceptualisation of 'metatheory' as 'systematic discourse about theory'. By expanding on this definition, I have explored a considerable number of implications and illustrated each of them with reference not only to IR, but also to parallel disciplines. Metatheory operates as a 'bridging' mechanism, constraining and enabling theoretical claims in many different ways. It often comprises an intermediate 'discursive layer' between one type of theoretical material and other types, within a given discipline. Secondly, the thesis also makes a thorough examination of IR views on metatheoretical research, weighing arguments both in favour and against it. A key finding is that, even from the perspective of each position, some of the arguments contradict each other. Moreover, most of the objections to metatheorising in our discipline are not of the 'strong' kind. Instead, contingent claims are made about negative effects of (too much) metatheory in IR. Such arguments must be assessed on a case-by-case basis. Thirdly, my study shows the impossibility of establishing a strong argument for the complete eradication of metatheory in a given discipline and, by extension, in IR. In order to make a radical claim against metatheory, one has to engage in metatheoretical argumentation and, therefore, the exercise is self-referentially

inconsistent. Since we cannot fully remove metatheory from IR, a better way to proceed is to look at some of the problems in this discursive layer aiming to improve the quality of metatheoretical scholarship. This defence of metatheory, advanced in parts one and two of the thesis, should not be taken as an immunisation of metatheorising. One of the accounts of the negative roles played by this discursive layer in the field are, indeed, pertinent to the discussion and should be taken seriously. My defence, then, is a qualified one: we cannot get rid of metatheory, but it can and should be improved.

Secondary contributions

Part three of the thesis illustrates and further develops some of the general claims with the use of concrete 'cases' speaking to the areas of International Political Theory (IPT), History of International Relations and International Political Economy (IPE). Here I also claim originality, albeit in more localised terms. The first illustration case, studying political theories of the early Roman Empire written in the 1st century, examines the metatheoretical 'drives' behind each of the three key formulations. It may be seen from the perspective of IR as a contribution to the 'history of ideas' in the field. The notion of a 'states system', refined as an attempt to define the Empire in that discussion, was subsequently reframed by 19th century historians and later absorbed by English school theorists. The English school connection is clear, but it needs refinement. Whether it accounts for the origin of mainstream IR conceptions of the 'international system' also remains to be seen. This is an avenue for further research which I have helped to establish. The second illustration case addresses a number of theoretical interpretations of the impact of the Peace of Westphalia on the formation of the early (modern) international system. The theories differ not only in their reading of Westphalia, or substantive theoretical claims, but also in the metatheoretical 'drives' behind them. Particularly relevant in this case is the relation between theory and history presupposed by each approach. The argument that different IR theories would, and do, provide distinct interpretations of the same event is almost trivial. However, the way in which I explore this fact opens up a wide range of research issues to the IR scholar. In the first part of the thesis, I have alluded to peculiar readings of Thomas Kuhn's 'paradigmism', with the mainstream version attempting to 'normalise' IR as a 'science' and move on from heavy theorising to problem-solving empirical research. A common postpositivist reading of Kuhn tends to ignore this 'imperial bid' for 'normal science' and highlights the distinction between 'paradigms', leading to the postulation of 'incommensurability'

between them. (y own comparison between different theories of order and change in world politics applied to the case of 7estphalia is a modest starting point for those interested in further understanding this 'battlefield' for competing conceptions of how to do research in historical and empirical IR. Are we comparing different views of the same ob9ect8 Is there a partial overlap at least8 +r are these theories spea!ing of 'different worlds' in an incommensurable way8 A metatheory in favour or against 'incommensurability' would have to be specified, and the application could draw upon some of the findings contained in my illustration case.

&alen together, the chapter on .>th century theories of the -oly Roman #mpire and the study on order, change and 7estphalia also suggest a possible line of in"uiry to the historically'minded researcher. +ne could assume 0 as many do 0 that part of what happened in the reconstitution of the #mpire via 7estphalia and later developments derives from the social construction of #urope, federalism, international society, and so on. *nder this assumption, a theoretical framewor! could be designed (or borrowed) to account for the impact of political thought on political institutions and practices. &his would facilitate the connection between the .>th century theories and the theoretically' informed history of order and change in #arly (odern international society, bearing in mind the "uestion of how political thin!ers of that time interpreted the #mpire and the changes they witnessed in its development. &his type of 'constitutive' study would, therefore, be able to draw on my findings, but the focus here has been limited to the metatheoretical configuration of the debates on the #mpire (in I3&) and 7estphalia (in IR theory and history). In this sense, I have provided a more localised original contribution in my comparison'and'contrast interpretation of the .>th century theories and in its presentation to an IR audience.

&he third illustration also connects to the other cases, but in a less direct way. 7hen they assess the impact of 7estphalia, theories of order and change in the international system often allude to notions of hierarchy (before 7estphalia) and anarchy of sovereign states (after 7estphalia). 7hile the historical accuracy of these claims can be, and has, been challenged on many accounts (including some of the points raised by the approaches analysed in this thesis), the concepts of hierarchy and anarchy also re"uire further e\$amination. I have compared and contrasted a number of theories of international hierarchy. In the process, I have e\$amined their coherence with metatheoretical principles advocated by their authors themselves. &his 'immanent' critical procedure was complemented by an 'overarching' interpretation provided by a

recent study on the 'Eurocentric' character of international theory. The case further illustrates some claims advanced in this thesis about metatheory as a constraining and enabling 'mechanism', but also takes a step further and critically examines the overarching framework of interpretation itself. I have deliberately selected a theory already addressed by the framework, and found its reading of that theory wanting. From this very punctual discussion, it becomes clear that 'detailed reading' techniques can be a way of improving the quality of metatheoretical research in IR. Metatheory deals with theoretical tests, so in-depth interpretation is recommended as a preliminary procedure. Besides critiquing this 'overarching framework' in the final chapter, I have attempted to expand its application to other theories of hierarchy, including a mainstream approach previously unaccounted for (dyadic hierarchy), a 'non-Western' theory (peripheral realism) and a theory assembled from scattered arguments in a branch of political economy (praxeology). The specific original contributions here are, first, the evaluation of the overarching framework of interpretation of 'Eurocentrism' in international theory, together with its expanded application to novel cases; and, secondly, my inclusion of peripheral realism and introduction of the praxeological view of the 'hegemonic cycle' to the IR and I3 audience.

Questions, answers and evaluation

When taken together, the primary and the secondary contributions of this thesis make it possible to answer the questions raised in the very beginning. What is the role of metatheorising? How does it operate in IR? In what ways does metatheory shape theoretical research in our discipline? The general answer, obtained in part one, is that metatheory operates, on the one hand, in function of the subject-matter addressed in metatheoretical discourse and, on the other, in different modes of inquiry. Subject-matter may be divided according to a combination of the focus of metatheory (Internal or External to the discipline) and the 'point of entry' into theory (whether Intellectual or Ontological). From the possible combinations, we obtain Internal: Intellectual, Internal: Ontological, External: Intellectual and External: Ontological metatheorising. This is the subject-matter side. When it comes to modes of metatheoretical inquiry, at least four key 'ways of metatheorising' can be identified: hermeneutical, evaluative, corrective, critical and historical. In this scheme, respectively, metatheory interprets theory, judges it according to certain standards, refines it, provides a social critique of theory and accounts for its historical formation. In short, metatheory is a discursive

'mechanism' that constrains and enables certain types of scholarship. The first part of the thesis unpacks and briefly illustrates each of these points. With comparison and contrast of a number of theoretical approaches on different issues, the third part highlights, in greater detail, the workings of metatheory as a discursive mechanism in IR. This is the general answer to the question.

The literature on metatheory in our discipline also indicates particular answers to the question, which I have collected, organised and examined in detail in chapters F and G. IR scholars critical of (too much) metatheorising have raised points against it. Some of these arguments are cast in terms of intrinsic features of metatheory, a strong type of claim that purports to eliminate it from the field. Alternatively, weaker and contingent claims are made. These involve the complexity that metatheorising brings to IR, problems related to the communication of metatheoretical knowledge (especially to students and non-specialists) and the politics of metatheory (either disciplinary or general). There are also those who favour metatheoretical research, although they certainly do not assume the state of the art could not in any way be improved. For this group of scholars, metatheory may lead, and in the past has led, to better theory. It also helps us understand the discipline in a systematic way. In addition to this, metatheory clarifies and interprets theoretical material. Finally, this type of research may be employed as a framework to critically understand the role of theory in the social world (and the impact of the social world on theorising). I have initially taken all these claims at face value, but then cross-referenced them and detected contradictions in the group against metatheory. I have also contrasted positive and negative claims and pointed out that many of the positive claims adequately respond to the negative arguments. The main point derived from such moves, though, was my refutation of the 'strong' claim for the elimination of metatheory from IR. For all the merits of the contingent negative views of metatheorising, which should be weighed case-by-case, ultimately the bid against metatheory in IR fails to establish the strong claim.

The third part of the thesis answers the questions in a more narrow way. The general point expressed in part one is that metatheory constrains and enables certain intellectual procedures. Part two detects an intuitive notion of the general point in IR perceptions about the role of metatheory in the discipline. Part three takes a step further and illustrates how this operates in a number of different cases. The in-depth analyses of theoretical texts dealing with IR, theoretically-informed history of IR, IR theory and IR not only addressed how metatheory functions as a 'mechanism' in these concrete

cases. They also presented themselves as examples of metatheoretical research, further indicating what metatheory can do for us. In this thesis, therefore, answers to the questions raised are provided at the following levels% conceptual, inductive, illustrative and exemplary.

The conceptual answers are analytical in character, and should be judged according to their internal coherence. For instance, instead of asking whether my definition of 'metatheory' is 'wrong', one should attempt to demonstrate that the analytical claims do not follow from the starting point. The inductive claims, in turn, have been raised with basis on an examination of the relevant literature in IR, and should accordingly be evaluated with reference to that literature. However, if my contribution organising all those scattered passages on metatheory is deemed useful, I will consider the inductive claims successful enough. Another type of answer provided to the research questions here is illustrative% it refers to my attempt to uncover the mechanism of metatheorising in my 'illustration cases'. My analysis at this level should be compared and contrasted to the material selected for each case. It is indeed the case that tests are open to interpretation, but I have sought to justify each step of my claims with exhaustive reference to the sources and some of their commentators. These references are open to criticism, but they have been provided. Finally, the exemplary answer to the research question is performative, and emerges as an implication of the very procedure of arguing this thesis. The study as a whole, and the third part in particular, may be regarded as an illustration in itself of what metatheory can do for us. It may be negatively viewed as not complete enough from a philosophical perspective, but I see that rather as a virtue. If the procedures in this thesis help us better understand how metatheorising works, I would declare it successful also from an exemplary perspective.

Final remark*

As far as I can tell, the arguments contained in this thesis are clear and balanced enough in their circumscription of roles for metatheoretical research in the academic discipline of International Relations. The claim here is not that all scholars should metatheorise. The claim, rather, is that, if they are engaged in 'systematic discourse about theory', then in fact they are engaged in metatheoretical argumentation. This includes any strong attempts to eliminate metatheory from the discipline via 'systematic discourse about metatheory', and for this reason I have maintained that it would be logically impossible

to fully eradicate metatheorising from IR. #ven though negative views of this type of research may have a point in contingent terms, metatheory will remain part of IR as an academic discipline. If not all IR theorists are metatheorists, those who are would do well in pursuing a better understanding and e\$ecution of their research specialisation.

Bibliography

- Acharya, A. (2000). Dialogue and Discovery: In Search of International Relations Theories Beyond the West. *Millennium Journal of International Studies*, 4(1), 90-107.
- Achinstein, S. (1965). Models, Analogies, and Theories. *Philosophy of Science*, 42(1), 1-19.
- Achinstein, S. (1963). *Concepts of Science: A Philosophical Analysis*. Baltimore: Johns Hopkins Press.
- Adler, J. (1997). Defining the Middle Ground: Constructivism in World Politics. *European Journal of International Relations*, 4(1), 15-32.
- Adler, J. (2005). Constructivism and International Relations. In J. A. A. Immons (Ed.), *Handbook of International Relations* (pp. 9-30). London: Sage.
- Adler, J., & David Held. (1995). *How to Read a Book: The Classic Guide to Intelligent Reading* (2nd ed.). New York: Simon & Schuster.
- Aller, R. (1997). *Rediscoveries and Reformulations: Humanistic Methodologies for International Studies*. Cambridge: Cambridge University Press.
- Almond, I. (2000). *History of Islam in Human Thought: From Pre-Islam to Post-Islam*. London: Routledge.
- Althusser, L. (1971). *Politics: A Rigorous Translation of Politics Methodically Set forth and Illustrated with Sacred and Profane Examples*. (trans. Barney, Ed.). Indianapolis: Liberty Fund.
- Armstrong, C. (1994). Why is there too much International Theory? *Australian Journal of Political Science*, 40(1), 1-15.
- Aron, R. (1964). *Peace and War: A Theory of International Relations*. Garden City: Doubleday.
- Aron, R. (1969). What is a Theory of International Relations? *Journal of International Affairs*, 43(1), 1-15.
- Ashley, R. (1990). Political Realism and Human Interest. *International Studies Quarterly*, 34(1), 1-20.
- Ashley, R. (1991). The Covert of Neorealism. *International Organization*, 45(1), 1-25.
- Ashley, R. (1992). The Covert of Neorealism. In R. A. Cochrane (Ed.), *Neorealism and its Critics* (pp. 1-25). New York: Columbia University Press.
- Ashley, R. (1994). The Powers of Anarchy: Theory, Sovereignty, and the Domestication of Global Life. In J. A. A. Immons (Ed.), *International Theory: Critical Investigations* (pp. 1-25). Basingstoke: Macmillan.

- Ashley, R. <. (1998). The Achievements of Post-structuralism. In Smith, <. Loth, D. (eds.), *International Theory, Positivism and Beyond* (pp. 56/0-56F). Cambridge University Press.
- Ashworth, E. <. (1995). Did the Realist/Idealist Debate Really Represent a Revisionist History of International Relations. *International Relations*, 4(1), 11-20.
- Ashworth, E. <. (1996). Realism and the spirit of Alfred Thayer Mahan. *European Journal of International Relations*, 2(1), 1-10.
- Attin, T. <. (1996). *The Global Political System*. London: Sage.
- Aydinli, N. D. (1997). Periphery and theorising for a truly Internationalised Discipline: pinning IR theory out of Anatolia. *Review of International Studies*, 23(4), 619-635.
- Bachman, R. <. (1996). *Directions in Economic Methodology*. London: Routledge.
- Bain, <. (1997). The Inter-paradigm Debate. In <. Eight D.A. R. Brown (eds.), *International Relations and the Theory of International Relations* (pp. 105-120). London: Pinter.
- Bartelson, <. (1994). *J Genealogy of Sovereignty*. Cambridge University Press.
- Bates, <. R., D. <. E. (1997a). Teaching and Learning International Relations and Epistemology in Political Science. *Politics*, 12(1), 1-10.
- Bates, <. R., D. <. E. (1997b). In Defence of Pluralism in the Teaching of International Relations and Epistemology. *Politics*, 12(1), 11-20.
- Bauman, <. (1997). *Thinking Sociologically*. London: Sage.
- Bayer, <. (1997). The (ad)woman and the Languages of the International Feminist Theory. *Tulsa Studies in Literature*, 42(1), 1-10.
- Bell, <. (1997). *Il Piccolo Stato nell'Età Moderna: Studi su un concetto della Politica Internazionale tra il 1600 e il 1800*. (ilan% #di?ioni *niversitarie 4aca.
- Bell, <. (1997). The Westphalian Model in Defining International Law: Challenging the Myth. *Australian Journal of Legal History*, 12(1), 1-10.
- Bell, <. A., (ed.), <. D. <. E. (1997). Correspondence: Start the Revolution without us. *International Security*, 21(1), 1-10.
- Bhabha, <. (1997). Mimicry and the Ambivalence of Colonial Discourse. *7th to 10th*, (Special Issue on Psychoanalysis), 1-10.
- Bhaskar, <. A. (1997). *The Possibility of Naturalism* (2nd ed.). London: Routledge.
- Bilgin, <. (1997). The Changing Past of International Relations. *Third World Quarterly*, 18(1), 1-10.
- Bilgin, <. (1997). *History of Diplomacy*. London: Reaktion.
- Bilgin, <. (1997). The New Hermeneutics, or Paradigms Hermeneutics Research Programmes in the History of Economics. *History of Political Economy*, 29(1), 1-10.
- Bilgin, <. (1997). *The Methodology of Economics: How Economists Explain* (Revised.). Cambridge University Press.

- Lochman, J. (1999). What is to be Done? The Science Question in International Relations. *International Theory*, 4(F), 600-619.
- Loth, J. (1999). Care not to know? International Relations & theory versus the future. In J. Loth & J. Smith (eds.), *International Relations Theory Today* (pp. 15-30). Cambridge: Polity.
- Lotz, E. (1999). *In Introduction to the Philosophy of Science*. Cambridge: Polity.
- Louch, C. (1999). *Political Theories of International Relations*. Oxford: Oxford University Press.
- Louch, C. (1999). Resurrecting Ruggie and capturing the Westphalian moment. *Review of International Studies*, 25(1), 1-15.
- Loud, J. (1999). *Esquissée d'une Théorie de la Pratique*. Paris: CNRS.
- Loy, J. A., Zordon, C., Aluani, J., Asimovs, J., J. A., J. A., Zabi, Q., (2000), et al. (1999). Teaching International Studies from a Regional Perspective: An International Symposium on Power, Health and Global Order: An International Debate for Africa. *International Studies Perspectives*, 1(F), 1-10.
- Lynn, B. (1999). *The Peloponnesian War: Athens' Sparta' and the Struggle for Greece*. New York: Oxford University Press.
- Lynn, J. (1999). Reconsidering Wendt's (1999) theory of International Realism with social constructivism. *Journal of International Relations and Development*, 2(1), 1-15.
- Lynn, J. (1999). *International Relations Theory: Deconstructive Approaches*. London: Routledge.
- Lynn, J. (1999). IR & theory in Britain: The New Realism. *Review of International Studies*, 25(1), 1-15.
- Lynn, J. (1999). Situating Critical Realism. *Millennium: Journal of International Studies*, 1(1), 1-15.
- Lynn, J. (1999). The Practice & Theory of Realism: Towards a Chronological International Political Theory. *Millennium: Journal of International Studies*, 1(1), 1-15.
- Lynn, J., Bardin, J., D Rengger, B. (1999). Introduction. *International Relations in Political Thought: Texts from the Ancient Greeks to the First World War* (pp. 1-10). Cambridge: Cambridge University Press.
- Lynn, J. A. (1999). *Positivism in Social Theory and Research*. New York: Oxford University Press.
- Lynn, J. (1999). International & theory: The case for a classical Approach. *World Politics*, 41(F), 1-15.
- Lynn, J. (1999). New Directions in the Study of International Relations. *International Studies*, 41(1), 1-15.
- Lynn, J. (1999). *The Anarchical Society: A Study of Order in World Politics*. New York: Columbia University Press.

- Fullbrook, D., D Watson, A. (eds.). (1990). *The Expansion of International Society*. Oxford: Clarendon Press.
- Lunge, C. (1995). *Scientific Research in The Search for System*. Berlin: Springer-Verlag.
- Lunge, C. (1995). A Program for the semantics of science. *Journal of Philosophical Logic*, 4(F:G), F.50-55.
- Lunge, C. (1995). *Finding philosophy in social science*. Yale University Press.
- Lunge, C. (1996a). *Philosophy of Science: From Problem to Theory*. New Brunswick: Transaction.
- Lunge, C. (1996b). *Philosophy of Science: From Explanation to Justification*. New Brunswick: Transaction.
- Lunge, C. (1999a). *Social science under debate: a philosophical perspective*. University of Toronto Press.
- Lunge, C. (1999b). *The Sociology of Philosophy Connection*. New Brunswick: Transaction.
- Lunge, C. (2005). How Does It Work? The Search for Explanatory Mechanisms. *Philosophy of the Social Sciences*, 35(5), 505-15.
- Lurchill, D., D Einikater, A. (2005). Introduction. In Lurchill, A., Einikater, R., Cevetani, 4. Connelly, C., 3aterson, 6. Reusmit, D. 4. & rue (eds.), *Theories of International Relations* (Frieded.).asingstone%3algrave (acmillan).
- Lutman, L., D -ansen, E. (2009). *The Evolution of International Security Studies*. Cambridge: Cambridge University Press.
- Lutman, L., D Eittle, R. (2011). *International Systems in "world history: Remaking the Study of International Relations*. Oxford: Oxford University Press.
- Lutman, L., D Eittle, R. (2011). Why International Relations has failed as an Intellectual Project and What to do About it. *Millennium: Journal of International Studies*, 19(1), 90-109.
- Lutman, L., D Eittle, R. (2009). Introduction to the 2009 Reissue. In A. Watson (ed.), *The Evolution of International Society: A Comparative-Historical Analysis* (Reissue., p. i-500). London: Routledge.
- Maldwell, L. 4. (1996). *Beyond Positivism: Economic Methodology in the Twentieth Century* (Revised.). London: Routledge.
- Mallahan, L. (2005). *Economics for Real People: An Introduction to the Austrian School* (5th ed.). Auburn: Mises Institute.
- Malvin, L. (1990). *Institutio Christianae Religionis*. Geneva: Eerdmans.
- Mampbell, C. (1990). *Writing Security: United States Foreign Policy and the Politics of Identity* (Revised.). (Manchester: Manchester University Press).
- Mampbell, C. (2005). Poststructuralism. In & Cunne, C. (ed.), *International Relations Theories: Discipline and Diversity* (5th ed., pp. 505-515). Oxford: Oxford University Press.
- Mardoso, L. (2011). Analyse (emaria (recordab\esde#no, aletto). *Tempo Social*, 4(1), 5-1055.

- Barney, J. (199A). &ranslatorPs Introduction. *Politica9 Jn J ridged Translation of Politics Methodically Set forth and Illustrated/ith Sacred and Profane Examples*.
- Barvalho, I. de, Eeira, -, D -obson, 4. (. (5/..). &he lig langs of IR% &he (yths that your &eachers still &ell) ou about. =G@ and .9.9. *Millennium9 Journal of International Studies*, A+(F), >FA0>A@.
- Bervo, A. E. (5//@). *InserOPo Internacional9 FormaOPo dos lonceitos Brasileiros*. co 3aulo% araiva.
- Bernoff, J. (5//5). cientific Realism as a (eta' &heory of International Relations. *International Studies Quarterly*, G*(5), .@905/>.
- Bernoff, J. (5//>a). *Theory and Metatheory in International Relations9 loncepts and lontending Jccounts*. lasingsto!e% 3algrave (acmillan.
- Bernoff, J. (5//>b). cientific Realism, 6ritical Realism, and International Relations &heory. *Millennium9 Journal of International Studies*, AI(5), G/50G/>.
- Bernoff, J. (5//@). International Relations, 3aleontology, and cientific 3rogress% 3arallels between Cemocratic 3eace studies and the (eteor Impact #Stinction -ypothesis. *International Studies Perspectives*, +(.), 9/09@.
- Bernoff, J. (5//9a). Cefending, oundations for International Relations &heory. *International Theory*, 4(F), G==0G>>.
- Bernoff, J. (5//9b). 6onventionalism as an Ade"uate lasis for 3olicy'Relevant IR &heory. *European Journal of International Relations*, 4I(.), .A>0.9G.
- Bhristensen, &. 4., D nyder, 4. (.99>). 3rogressive Research on Cegenerate Alliances. *American Political Science Revie/*, +4(G), 9.90955.
- Biprut, 4. H. (5//). Introduction%(a'ling ense of International Relations &heory in 2lobal 6onte\$. In 4. H. Biprut (#d.), *The Jrt of the Feud9 Reconceptuali?ing International Relations* (p. \$i0\$\$vii). 7estport% 3raeger.
- Blar!, I. (.9@9). *The >ierarchy of States9 Reform and Resistance in the International 7rder*. 6ambridge% 6ambridge *niversity 3ress.
- Blar!, I. (5//9). &owards an #nglish hool &heory of -egemony. *European Journal of International Relations*, 4I(5), 5/F055@.
- Boats, A. 7. (.9=9). Is there a L tructure of cientific RevolutionsM in #conomics8 Qy%los, (((5), 5@9059=.
- Bochran, (. (.999). *Dormative Theory in International Relations9 J Pragmatic Jpproach*. 6ambridge% 6ambridge *niversity 3ress.
- Bollins, R. (.9@=). Is .9@/s ociology in the Coldrums8 *American Journal of Sociology*, +4(=), .FF=0.FAA.
- Bo\$, R. 7. (.9@.). ocial ,orces, tates and 7orld +rders% leyond International Relations &heory. *Millennium9 Journal of International Studies*, 4)(5), .5=0.AA.
- Bo\$, R. 7. (5//@). &he 3oint is not 9ust to #\$plain the 7orld but to 6hange it. In 6. Reus' mit D C. nidal (#ds.), *The 7xford >and oo% of International Relations* (pp. @G09F). +\$ford% +\$ford *niversity 3ress.
- Bo?ette, (. (5//@). Reclaiming the 6ritical Cimension of Realism%-ans 4. (orgenthau on the #thics of holarship. *Revie/ of International Studies*, AG(.), A05>.

- Aston, C. (1999). The New Revolution in Political Science. *American Political Science Review*, 93(4), 1001-1010.
- Adkins, L., & Haughan, B. (1999). Introduction. In L. Adkins & B. Haughan (Eds.), *Critical Theorists and International Relations* (pp. 1-10). London: Routledge.
- Einstein, A. (1956). *The Einstein Reader*. New York: Citadel Press.
- Farber, C. (1999). *Exploring Federalism*. Tuscaloosa: University of Alabama Press.
- Farber, C. (1999). *Covenant and Commonwealth: From Christian Separation through the Protestant Reformation*. New Brunswick: Transaction.
- Idman, M., & Idman, C. (1999). Realism and Neorealism: A Reply to Hasenue. *American Political Science Review*, 93(4), 950-955.
- Idman, M., & Idman, C. (1995). How not to be Realist: Appraising Progress in IR Research. *International Studies Quarterly*, 39(5), 1051-1065.
- Idman, M., & Idman, C. (1995). Introduction: Appraising Progress in International Relations Theory. In M. Idman & C. Idman (Eds.), *Progress in International Relations Theory: Appraising the Field* (pp. 1-105). Cambridge, MA: MIT Press.
- Ishtain, L. (1999). International Politics and Political Theory. In L. Ishtain & L. Ishtain (Eds.), *International Relations Theory Today* (pp. 1-105). Cambridge: Polity.
- Ishtain, L. (1999). *Explaining Social Behavior: More Dots and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
- Jopp, R. (1999). The English School on the Frontiers of International Society: A Hermeneutic Reflection. *Review of International Studies*, 25(4), 601-610.
- Kristine, E. (1999). Normative International Relations Theory. In E. Kristine & L. Cunne, C. (Eds.), *International Relations Theories: Discipline and Diversity* (5th ed., pp. 1-105). Oxford: Oxford University Press.
- Schle, M., & Aiguashca, L. (1999). Rethinking Globalised Resistance: Feminist Activism and Critical Theorising in International Relations. *British Journal of Politics and International Relations*, 5(5), 501-510.
- ScudJ, M. (1999). *Foreign Policy Theory in Menem's Argentina*. Gainesville: University Press of Florida.
- ScudJ, M. (1999). Peripheral Realism: An Argentine Theory Building Experience, 1999-1999. In L. Ishtain & L. Ishtain (Eds.), *Concepts, Histories and Theories of International Relations for the 21st Century: Regional and National Approaches* (pp. 105-110). Brasilia: IIRL.
- ScudJ, M. (1999). State personhood, reality or fiction? Buenos Aires.
- Ulau, R. (1999). Theories of Federalism under the Holy Roman Empire. *American Political Science Review*, 93(4), 1001-1010.
- Ulau, R. (1999). *The Behavioral Persuasion in Politics*. New York: Random House.
- Yeraband, S. (1999). *Against Method* (4th ed.). London: Herso.

- ierle, C. (5//>). Constructivism. In & Cunne, C. (urli, D. mith (#ds.), *International Relations Theories9 <iscipline and <iversity* (pp. .==0.@G). +\$ford% +\$ford *niversity 3ress.
- ingfeld, C. E. (5//F). (etasyntesis% &he tate of the Art ' so far. *Qualitative >ealth Research*, 4A(>), @9F09/G.
- is!e, C. 7., D hweder, R. A. (#ds.). (.9@=). *Metatheory in Social Science9 Pluralisms and Su2ectivities*. 6hicago% *niversity of 6hicago 3ress.
- lec!, E. (.9>9). *Henesis and <evelopment of a Scientific Fact*. (& 4. &renn D R. <. (erton, #ds.). 6hicago% &he *niversity of 6hicago 3ress.
- lyvb9erg, I. (5//.). *Ma%ing Social Science Matter9 "hy Social In\$uiry Fails and >o/ it can Succeed.Jgain*. 6ambridge% 6ambridge *niversity 3ress.
- o\$, 2. (.99>). *Reason and Reality in the Methodologies of Economics9 Jn Introduction*. 6heltenham% #dward #lgar.
- o\$, 7. & R. (5//.). International Relations &heory and Areas of 6hoice in ,oreign 3olicy. In B. 2uilhot (#d.), *The Invention of International Relations Theory9 Realism' the Roc%efeller Foundation' and the 4+IG lonference on Theory* (pp. 5>F05>A). Bew)or!% 6olumbia *niversity 3ress.
- ran!el, 4. (.9@.). 6onventional and &heorising Ciplomats% A 6riti"ue. *International Jffairs*, IM(G), AF>0AG@.
- reire, E. 2. (5//=a). Beo'realismo e 6onduta Incoerente com a #strutura do istema Internacional. *Fronteira*, I(9), 9>0.5A.
- reire, E. 2. (5//=b). *Js MVltiplas 6o?es de Qenneth "alt?9 Um Estudo de Segunda 7rdem em RelaOWes Internacionais*. 3ontifdcia *niversidade 6atalica de (inas 2erais.
- reire, E. 2. (5//=c, April). olidco na (ultidco% 6onsiderab\es sobre as Relab\es Internacionais na #uropa (edieval. 7 <e atedouro.
- reire, E. 2. (5//@). *The Use of Imre @a%atos.s Philosophy of Science in International Relations9 Some Pro lems of Interpretation*. *niversity of #\$eter.
- reire, E. 2. (5//9). Agents, *niversality and tructural 6hange% (eta'&heoretical Remar!s and an Agenda for ocial Research. >uSS PH lonference (p. .@).
- reire, E. 2. (5//.). 7hat can we Eearn from Althusius8 +ntology, 3luralism and 2lobalisation. *4+th Hlo alJ/areness Society lonference*. <ra!ow.
- reire, E. 2. (5//.). &he Cawn of tates' ystem as a 6oncept% Althusius, 3ufendorf and Eeibni? on the -oly Roman #mpire. (ontrJal% l A Annual 6onference. Retrieved from [http%:::www.allacademic.com:meta:pA/.>@.einde\\$.html](http%:::www.allacademic.com:meta:pA/.>@.einde$.html)
- reire, E. 2. (5//5). (artin 7ight on 7ar% &owards a letter *nderstanding of the L#nigma.MISJI Ard Jnnual lonvention (p. .@). an Ciego.
- reire, E. 2., D <oivisto, C. (5//5). International Relations as a ocial cience. In C. Armstrong (#d.), *7xford Bi liographies 7nline9 International Relations*. Bew)or!% +\$ford *niversity 3ress.
- reire, 3. (5//=). *Pedagogy of the 7ppressed* (F/th anniv.). Bew)or!% 6ontinuum.

- rieden, A. A., D Eale, C. A. (5//A). International Relations as a social science% Rigor and Relevance. *Jnnals JJPSS*, *)), .F=0.A=.
- riedman, C. (.9AF). *Essays in Positive Economics*. Chicago% *niversity of Chicago Press.
- rigg, R., D -artmann, . (5//9). Models in science. *Stanford Encyclopedia of Philosophy*. Center for the Study of Language and Information. Retrieved from <http://plato.stanford.edu/archives/sum5//9:entries/models/science>
- rit?, A. (5//A). Internationalism and the Promise of science. In C. Eong D 1. 6. Schmidt (#ds.), *Imperialism and Internationalism in the <iscipline of International Relations* (pp. .G.O.A@). Albany% *B) Press.
- rost, L.'3. (.99>). Resurrecting a Beglected &heorist% the Philosophical Foundations of Raymond AronPs &heory of international Relations. *Review of International Studies*, A(5), .GF0.==.
- rost, C. (.99=). *Ethics in International Relations9J l onstitutive Jpproach*. Cambridge% Cambridge *niversity Press.
- rJmont, G. (.99=). Leibni?, le (odfle #uropJen contre la (onarchie T la ,ranbaise. In -. (Jchoulan D 4. Gornette (#ds.), @.Xtat llassi\$ue' 4*I(04M4I (pp. .=.0.@/). Paris% 4. Hrin.
- rJmont, G. (5//F). *SingularitL9 Individus et Relations dans le SystYme de @ei ni?*. Paris% 4. Hrin.
- urlong, S., D arsh, C. (5//.). A !in not a weater% +ntology and #pistemology in Political science. In C. arsh D 2. to!er (#ds.), *Theory and Methods in Political Science* (Frd ed., pp. .@G05..).asingsto!e% 3algrave (acmillan.
- abriel, A. C. (.99G). *orldvie/s and Theories of International Relations*.asingsto!e% (acmillan.
- allie, T. L. (.9>@). *Philosophers of Peace and "ar9 Qant' llause/it?' Marx' Engels and Tolstoy*. Cambridge% Cambridge *niversity Press.
- teert?, G. (.9@F). *@ocal Qno/ledge9 Further Essays in Interpretive Jnthropology*. New York% Basic Books.
- teorge, A. (.9@9). International Relations and the Search for &hin!ing pace% Another View of the &hird Debate. *International Studies Quarterly*, AA(F), 5=905>9.
- teorge, A. (.995). Some Thoughts on the Liveness of #veryday Life in Australian International Relations% &heory and Practice. *Australian Journal of Political Science*, M(.), F.OAG.
- teorge, A. (.99G). *<iscourses of Hlo al Politics9J lritical &Re-Introduction to International Relations*. Boulder% Eynne Rienner.
- Tiddens, A. (.9>9). *entral Problems in Social Theory9 Jction' Structure' and l contradiction in Social Jnalysis*. Berkeley% *niversity of California Press.
- Tiere, R. B. (.9@@). *Explaining Science9J lognitive Jpproach*. Chicago% *niversity of Chicago Press.
- Tiere, R. B. (5//G). How Models Are Used to Represent Reality. *Philosophy of Science*, M4(A), >G50>A5.

- 2u??ini, . (5/.). Imposing coherence% the central role of practice in ,riedrich
<ratochwilPs theorising of politics, international relations and science. *Journal of
International Relations and <velopment*, 44(F), F/.0F55.
- 2u??ini, . (5/.). ecuriti?ation as a 6ausal (echanism. *Security <ialogue*, G((G),
F590FG..
- 2u??ini, ., D Eeander, A. (#ds.). (5//=). *Ionstructivism and International Relations9
Jlexander "endt and his Iritics*. Eondon% Routledge.
- age, 4. (.9>5). *Techni\$ues and Pro lems of Theory Ionstruction in Sociology*. Bew
)or!% 4ohn 7iley D ons.
- all, I. (5//=). *The International Thought of Martin "ight*. Bew)or!% 3algrave.
- alliday, ,. (.99G). *Rethin%ing International Relations*. 1asingsto!e% 3algrave.
- alliday, ,. (.99A). International Relations and Its Ciscontents. *International Jffairs*,
M4(G), >FF0>G=.
- alliday, ,. (.99=). &he ,uture of International Relations% ,ears and -opes. In . mith,
<. 1ooth, D (. Qalews!i (#ds.), *International Theory9 Positivism and Beyond* (pp.
F.@0F5>). 6ambridge% 6ambridge *niversity 3ress.
- alliday, ,., D Rosenberg, 4. (.99@). Interview with <en 7alt?. *Revie/ of International
Studies*, (G(F), F>.0F@=.
- ands, C. 7. (.99F). 3opper and Ea!atos in #conomic (ethodology. In *. (I!i, 1.
2ustafsson, D 6. <nudsen (#ds.), *Rationality' Institutions and Economic
Methodology* (pp. =.0>A). Eondon% Routledge.
- aslam, 4. (5//=). *J Decessidade L a Maior 6irtude9 7 Pensamento Realista nas
RelaOWes Internacionais*. co 3aulo% (artins,ontes.
- ausman, C. (. (5//@). Introduction. In C. (. -ausman (#d.), *The Philosophy of
Economics9 Jn Jnthology* (Frd ed., pp. .0F@). 6ambridge% 6ambridge *niversity
3ress.
- avercroft, 4. (5/.5). 7as 7estphalia Lall thatM8 -obbes, 1ellarmine, and the Borm of
Bon'Intervention. *Hlo al Ionstitutionalism*, 4(.), .5/0.G/.
- ay, 6. (5//5). *Political Jnalysis9 J Iritical Introduction*. 1asingsto!e% 3algrave.
- ay, 6. (5//>). Coes +ntology &rump #pistemology8 Botes on the Cirectional
Cependence of +ntology and #pistemology in 3olitical Analysis. *Politics*, (M(5),
..A0..@.
- edstrRm, 3., D Jli!os!i, 3. (5/.). 6ausal (echanisms in the ocial ciences. *Jnnual
Revie/ of Sociology*, A*, G90=>.
- eeren, A. -. E. (.@FG). *Manual of the >istory of the Political System of Europe and its
Iolonies' from its Formation at the Ilose of the Fifteenth Ientury to its Re0
Esta lishment upon the Fall of Dapoleon*. +\$ford% C. A. &alboys.
- eld, C. (.99A). <emocracy and the Hlo al 7rder. tanford% tanford *niversity 3ress.
- empel, 6. 2. (.9=A). *Jspects of Scientific Explanation9 Jnd 7ther Essays in the
Philosophy of Science*. Bew)or!% ,ree 3ress.

- endry, R. J., D. S. (1995). *How to do things with theories: An Interactive View of Language and Models in Science*. In J. R. S. (ed.), *The Language of Philosophy: Essays Dedicated to J. R. S.* (pp. 1-10). Amsterdam: Rodopi.
- er, J. (1999). *International Politics in the Atomic Age*. New York: Columbia University Press.
- er, J. (1999). Relevancies and Irrelevancies in the Study of International Relations. *Polity, G*(.), 5A0G.
- ess, C. J. (1999). *Science Studies: An Advanced Introduction*. New York: New York University Press.
- obson, J. (1995). *The Eurocentric Conception of World Politics: Western International Theory* (4th ed.). Cambridge: Cambridge University Press.
- offmann, J. (1999). An American Social Science: International Relations. (4. Cer Cerian, J. (ed.), *Aedalus*, 4)(F), G.0=.
- olden, J. (1995). The Ontological and Epistemological Disciplinary History and the Discourse about IR Discourse. *Review of International Studies*, 21(5), 5AF0 5>.
- ollis, J. (1995). *The Philosophy of Social Science: An Introduction* (Revised ed.). Cambridge: Cambridge University Press.
- ollis, J., D. Smith, J. (1999). *Explaining and Understanding International Relations*. Oxford: Oxford University Press.
- ollis, J., D. Smith, J. (1999). Beware of Errors: Structure and Action in International Relations. *Review of International Studies*, 25(G), F9F0G./.
- ollis, J., D. Smith, J. (1999). A Response: Why Epistemology Matters in International Theory. *Review of International Studies*, 25(1), ...0..=.
- olsti, J. (1999). *The Dividing Discipline: Hegemony and Diversity in International Theory*. London: Allen D. Stein.
- olsti, J. (1999). Error, Error on the Fall, Which Are the Fairest Theories of All? *International Studies Quarterly*, 43(F), 5AA05=.
- olsti, J. (1999). Scholarship in an Era of Anxiety: The Study of International Politics during the Cold War. *Review of International Studies*, 25(A), >0G=.
- olsti, J. (1999). Obstacles to Understanding in International Relations. In J. H. G. (ed.), *The Art of the Feud: Reconceptualizing International Relations* (pp. 5>0G=). Westport: Praeger.
- oover, J. R. (1999). *The Elements of Social Scientific Thinking* (2nd ed.). New York: Martinus Press.
- opf, J. (1999). The Promise of Constructivism in International Relations Theory. *International Security*, 24(1), >.05//.
- oppe, J. (1999). (Aristotle and Aristotle's Analysis. *Journal of American Studies*, 33(5), >909F.

- Joseph, D., D. Light, G. (eds.). (5/1/). *Scientific Realism and International Relations*. London: Routledge.
- Kousser, L. de. (1997). The Nature of Political Science. *American Political Science Review*, 92(1), 1-10.
- Kupila, A. (5/1/). Knowing Europe: (ethnography and methodology in studies. In C. Gini D. A. Bourne (eds.), *Palgrave Advances in European Union Studies* (pp. 5/90-5/95). London: Palgrave.
- Kurzman, C. (5/1/). Our Levels and a Discipline. In C. Kurzman & C. Kurzman (eds.), *Constructing International Relations: The Next Generation* (pp. 1-10). London: Routledge.
- Kurzman, C. (5/1/). *International Relations Theory: A Critical Introduction*. London: Routledge.
- Kurzman, C. (1997). Teaching International Relations in a Changing World: Our Approaches. *PS: Political Science and Politics*, 1(1), 1-10.
- Kurzman, C. (5/1/). The Theoretical Roots of Hierarchy in International Relations. *Australian Journal of International Affairs*, 1(1), 1-10.
- Kurzman, C. A. (1997). The New Great Debate: Realism vs. Science in International Relations. *World Politics*, 40(1), 1-10.
- Kurzman, C. A. (5/1/). *System and Process in International Politics* (5th ed.). London: Routledge.
- Kurzman, C. A., D. Light, R. (5/1/). Eclectic Theorizing in the Study and Practice of International Relations. In Christian Reus-Mietz & C. Kurzman (eds.), *The Oxford Handbook of International Relations* (pp. 1/90-1/95). Oxford: Oxford University Press.
- Kurzman, C. A., Eittle, R., D. Holford, G. (eds.). (5/1/). *Balance of Power in World History*. London: Routledge.
- Kurzman, C. (5/1/). Testphalian Eurocentrism in International Relations Theory. *International Studies Review*, 4(1), 1-10.
- Kurzman, R., D. Kury, A. (1995). *Social Theory as Science* (5th ed.). London: Routledge and Kegan Paul.
- Kurzman, C. (5/1/). *Beyond the Anarchical Society: Hobbes' Colonialism and Order in World Politics*. Cambridge: Cambridge University Press.
- Kurzman, C. (5/1/). *International Political Thought: A Historical Introduction*. Cambridge: Polity.
- Kurzman, A. (5/1/). *A History of Western Philosophy* (6 volumes). Oxford: Oxford University Press.
- Kurzman, R. +. (1997). Realism, Neorealism and the Study of World Politics. In R. +. Kurzman (ed.), *Neorealism and its Critics* (pp. 1-10). New York: Columbia University Press.
- Kurzman, R. +. (1997). International Institutions: Two Approaches. In C. Ratochwil & C. Ratochwil (eds.), *International Studies Quarterly*, 41(1), 1-10.

- Keohane, R. O. (1990). International Relations and the contributions of a feminist standpoint. *Millennium Journal of International Studies*, 4(5), 5GA05AF.
- Keohane, R. O., & D. Byrd (1990). *Power and Interdependence: World Politics in Transition* (2nd ed.). Toronto: E. J. Brill, 1990.
- King, B., Keohane, R. O., & D. Herbert (1998). *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton: Princeton University Press.
- Klotz, A., & D. Eynon (1990). *Strategies for Research in Constructivist International Relations*. Armonk: C. J. Harper.
- Knutson, E. (1990). *History of International Relations Theory* (5th ed.). Manchester: Manchester University Press.
- Kornprobst, C. (1990). International Relations as Rhetorical Discipline: Toward (Re)Bewing -orientations. *International Studies Review*, 44(1), 1-10.
- Krasner, S. C. (1990). Testphalia and all that. In S. D. Stein & R. O. Keohane (eds.), *Ideas and Foreign Policy: Beliefs, Institutions, and Political Change* (pp. 5FA05=G). Ithaca: Cornell University Press.
- Kratochwil, F. H. (1990). Theory and Political Practice: Reflections on Theory Building in International Relations. In S. J. D. E. R. Rui (eds.), *Principled World Politics: The Challenge of Dormative International Relations* (pp. A/0=G). Eanham: Rowman & Littlefield.
- Kratochwil, F. H. (1990a). If false promises and good bets: a plea for a pragmatic approach to theory building (the Krasner lecture). *Journal of International Relations and Development*, 4(1), 10A. doi:10.1016/j.palgrave.1990.101010
- Kratochwil, F. H. (1990b). Evidence, Inference, and Truth as Problems of Theory Building in the Social Sciences. In R. B. Eebow D. L. Eichbach (eds.), *Theory and Evidence in Comparative Politics and International Relations* (pp. 5A0AG). Iasingstone: Palgrave Macmillan.
- Kratochwil, F. H. (1990). *The Possibilities of Politics: Inquiries into the Genesis and Transformation of International Relations*. London: Routledge.
- Krauthammer, B. (1990). The bipolar moment. *Foreign Affairs*, 69(1), 5F0FF.
- Kuhn, T. S. (1990). *The Structure of Scientific Revolutions* (2nd ed.). Chicago: University of Chicago Press.
- Kurki, C. (1990). Causes of a Cived Discipline: Rethinking the Concept of Cause in International Relations Theory. *Review of International Studies*, 16(5), 1905-19.
- Kurki, C. (1990). *Lausation in International Relations: Reclaiming Lausal Analysis*. Cambridge: Cambridge University Press.
- Kurki, C. (1990). The Politics of the Philosophy of Science. *International Theory*, 4(F), GG/0GAG.
- Kurki, C., & D. Wright (1990). International Relations and Social Science. In S. J. D. E. R. Rui (eds.), *Principled World Politics: The Challenge of Dormative International Relations* (pp. F0FF). Oxford: Oxford University Press.
- Kurki, C., & D. Wright (1990). International Relations and Social Science. In S. J. D. E. R. Rui (eds.), *Principled World Politics: The Challenge of Dormative International Relations* (5th ed., pp. G0FA). Oxford: Oxford University Press.

- <urth, 4. (.99@). Inside the 6ave% &he 1anality of I.R. studies. *The Dational Interest*, 1A, 590G/.
- Ea!atos, I. (.9>/). ,alsification and the (ethodology of scientific Research 3rogrammes. In I. Ea!atos D A. (usgrave (#ds.), *Iriticism and the Hro/th of Qno/ledge* (pp. 9.0.9=). 6ambridge% 6ambridge *niversity 3ress.
- Ea!atos, I. (.9>.). -istory of science and its Rational Reconstructions. In R. 6. 1uc! D R. 6. 6ohen (#ds.), *Boston Studies in the Philosophy of Science'v!*, (pp. 9.0.F=). Cordrecht% C. Reidel.
- Ea!atos, I. (.9>@). *The Methodology of Scientific Research Programmes9 Philosophical Papers 6olume 4.* (4. 7orrall D 2. 6urrie, #ds.). 6ambridge% 6ambridge *niversity 3ress.
- Ea!e, C. A. (.99=). Anarchy, -ierarchy, and the Hariety of International Relations. *International 7rgani?ation, I*(.), .0FF.
- Ea!e, C. A. (5//>). #scape from the tate of Bature% Authority and -ierarchy in 7orld 3olitics. *International Security, A*((.), G>0>9.
- Ea!e, C. A. (5//9). >ierarchy in *International Relations*. Ithaca% 6ornell *niversity 3ress.
- Ea!e, C. A. (5//.). Rightful Rules% Authority, +rder, and the ,oundations of 2lobal 2overnance. *International Studies Quarterly, IG, A*@>0=F.
- Ea!e, C. A. (5//.). 7hy LismsM are #vil% &heory, #pistemology, and Academic ects as Impediments to *nderstanding and 3rogress. *International Studies Quarterly, II*(5), G=A0G@/.
- Eapid,). (.9@9). &he &hird Cebate% +n the 3rospects of International &heory in a 3ost' 3ositivist #ra. *International Studies Quarterly, AA*(F), 5FA05AG.
- Eatsis, . (#d.). (.9>=). *Method and Jppraisal in Economics*. 6ambridge% 6ambridge *niversity 3ress.
- Eawrence, A. (5//>). Imperial 3eace or Imperial (ethod8 !eptical In"uiries into Ambiguous #vidence for the LCemocratic 3eace.M In R. B. Eebow D (. I. Eichbach(#ds.), *Theory and Evidence in Iomparative Politics and International Relations* (pp. .99055=). 1asingsto!e% 3algrave (acmillan.
- Eawson, &. (.999). 7hat has Realism 2ot to Co with It8 *Economics and Philosophy, 4I*(5), 5=905@5.
- Eayne, 6. (.99F). &he *nipolar Illusion% 7hy Bew 2reat 3owers 7ill Rise. *International Security, 4M*(G), A0A..
- Eebow, R. B. (5//>). 7hat can we <now8 -ow do we <now8 In R. B. Eebow D (. I. Eichbach(#ds.), *Theory and Evidence in Iomparative Politics and International Relations* (pp. .055). 1asingsto!e% 3algrave (acmillan.
- Eebow, R. B., D Eichbach, (. I. (#ds.). (5//>). *Theory and Evidence in Iomparative Politics and International Relations*. 1asingsto!e% 3algrave (acmillan.
- Eeibni?, 2. 7. (.9@G). Ce 4ure uprematus ac Eegationis. In 3. Ritter D E. <nabe (#ds.), *SUmliche Schriften und Briefe*. 1erlin% A!ademie Herlag.

- Einbner, J. (1997). *Caesar's Commentaries on the Gallic Wars*. In J. E. Riley (ed.), *Caesar's Commentaries* (5th ed., pp. 1-105). Cambridge: Cambridge University Press.
- Eieder, R. (1997). *Theory and World Politics*. London: Allen & Unwin.
- Einphart, A. (1997a). International Relations & Theory: Great Debates and Lesser Debates. *International Social Science Journal*, 45(1), 1-10.
- Einphart, A. (1997b). The Structure of the Theoretical Revolution in International Relations. *International Studies Quarterly*, 41(1), 1-10.
- Einlater, A. (1995). *Men and Ideas in the Theory of International Relations*. London: Macmillan.
- Einlater, A. (1995). The Question of the Best Stage in International Relations & Theory: A Critical & Theoretical Point of View. *Millennium Journal of International Studies*, 4(1), 1-10.
- Einlater, A. (1997). The Achievements of Critical Theory. In J. Smith, C. Booth, D. C. Q. (eds.), *International Theory: Positivism and Beyond* (pp. 1-10). Cambridge: Cambridge University Press.
- Einlater, A. (1997). *Critical Theory and World Politics: Ideology, Sovereignty and Humanity*. London: Routledge.
- Einlater, A. (1997). The English School Conception of International Society: Reflections on Western and non-Western Perspectives. *Review of International Studies*, 23(1), 1-10.
- Einlater, A., & Suganami, H. (1997). *The English School of International Relations: A Contemporary Reassessment*. Cambridge: Cambridge University Press.
- Euomah, J. (1997). Political Ideology, Anthropomorphism, and Personhood of the State & the Religion of IR. *International Political Sociology*, 4(1), 1-10.
- Eyas, J. (1997). *Peter's Inch & Eddington's Acumen*.
- Eyotard, J. (1997). *La Condition Postmoderne: Rapport sur le Savoir*. Paris: Editions de l'Inuit.
- ERwy, J. (1997). *Paysages de la Littérature: Introduction à une Sociologie Critique de la Littérature*. Paris: Anthropos.
- (ac)ie, J. E. (1997). *The Element of the Universe: A Study of Causation*. Oxford: Oxford University Press.
- (alinia), C., & al., A., Peterson, J., & Tierney, J. (1997). International Relations in the Academy. *International Studies Quarterly*, 41(1), 1-10.
- (anning, J. A. (1997). *The University Teaching of Social Sciences: International Relations*. Paris.
- (arsh, C. (1997). Theoretical Issues. In C. (arsh D. J. (ed.), *Theory and Methods in Political Science* (Frd ed., pp. 1-10). Fasingstone & Algrave (acmillan).
- (arsh, C., & Tierney, J. (1997). *Theory and Methods in Political Science* (Frd ed.). Fasingstone & Algrave.

- (orgenthau, -. 4. (.9==). &he 3urpose of 3olitical cience. In 4. 6. 6harlesworth (#d.), *J<esign for Political Science9 Scope' 7 2ectives and Methods* (pp. =F0>9). 3hiladelphia% American Academy of 3olitical and ocial cience.
- (orrison, (. (5//>). 7here -ave all the &heories 2one8 *Philosophy of Science*, MG(5), .9A055@.
- (urdoc!, 2. (5//G). *Beyond 1alvin9 The Intellectual' Political and 1ultural "orld of Europe's Reformed 1hurches*. 1asingsto!e% 3algrave(acmillan.
- (I!i, *. (.99@). Realism. In 4. 1. Cavis, C. 7. -ands, D *. (I!i (#ds.), *The >and oo% of Economic Methodology* (pp. G/G0G/9). 6heltenham% #dward #lgar.
- Bagel, #. (.9=). *The Structure of Science9 Pro lems in the @ogic of Scientific Explanation*. Eondon% Routledge and <egan 3aul.
- Bau, -. R. (5/..). Bo Alternative to Llsms.M *International Studies Quarterly*, II(5), G@>0G9..
- Beal, ,. 7., D-amlett, 1. C. (.9=9). &he Bever'Bever Eand of International Relations. *International Studies Quarterly*, 4A(F), 5@.0F/A.
- Bel, 3., D (c2owan, 3. 4. (#ds.). (.999). *Po/er' "ealth and Hlo al 7rder9 Jn International Relations Text oo% for Jfrica* (.st ed.). 6ape &own% *niversity of 6ape &own 3ress.
- Beufeld, (. (.99.). &he Refle\$ive &urn in International Relations &heory. Cownsview.
- Beufeld, (. (.99G). Refle\$ivity and International Relations &heory. In 6. &. 9olander D 7. . 6o\$ (#ds.), *Beyond Positivism9 Iritical Reflections on International Relations* (pp. ..0FA). Eondon% Eynne Rienner.
- Beufeld, (. (.99A). *The Restructuring of International Relations Theory*. 6ambridge% 6ambridge *niversity 3ress.
- Bicholson, (. (.995). *Rationality and the Jnalysis of International Ionflict*. 6ambridge% 6ambridge *niversity 3ress.
- Bicholson, (. (.99=a). &he 6ontinued ignificance of 3ositivism8 In . mith, <. 1ooth, D (. Qalewsl!i (#ds.), *International Theory9 Positivism and Beyond* (pp. .5@0.GA). 6ambridge% 6ambridge *niversity 3ress.
- Bicholson, (. (.99=b). *1auses and Ionse\$uences in International Relations9 J 1onceptual Study*. Eondon% ,rances 3inter.
- Bicholson, (. (5///). 7hatPs the use of International Relations8 *Revie/ of International Studies*, (*(5), .@F0.9@.doi%./..!>:/5=/5./A///.@F5
- Biebuhr, R. (5/..). &he (oral Issue in International Relations. In B. 2uilhot (#d.), *The Invention of International Relations Theory9 Realism' the Roc%efeller Foundation' and the 4+IG Ionference on Theory* (pp. 5=905>5). Bew)or!% 6olumbia *niversity 3ress.
- Biiniluoto, 1. (5/..). cientific 3rogress. *Stanford Encyclopedia of Philosophy*. Retrieved from [http%::plato.stanford.edu/archives/sum5/...entries/scientific/progress](http://plato.stanford.edu/archives/sum5/...entries/scientific/progress):
- Bi9man, 4. #. (5//G). Eeibni?Ps &heory of Relative overeignty and International Eegal 3ersonality% 4ustice and tability or the East 2reat Cefence of the -oly Roman #mpire. Bew)or!.

- Bye, 4. . 4. (5//@). Bridging the Gap between Theory and Policy. *Political Psychology*, (+)(G), A9F0=F.
- +berheim, #. (5//=). *Feyerabend's Philosophy*. Berlin% Walter de Gruyter.
- +nuf, B. 2. (.9@9). "World of our Making: Rules and Rule in Social Theory and International Relations. Columbia% University of South Carolina Press.
- +siander, A. (.99G). *The States System of Europe' 4*G)04++) Peacemaking and the Conditions of International Stability*. +\$ford% +\$ford University Press.
- +siander, A. (5//.). Sovereignty, International Relations, and the Westphalian Myth. *International Organization*, II(5), 5A.05@>.
- +ssewaarde, C. R. R. (5//>). Three Rival Versions of Political Theory: Althusser and the Concept of Sphere Sovereignty. *The Monist*, +(.), ./=0.5A.
- +uthwaite, 7. (.9@>). *De/Philosophies of Social Science: Realism's Hermeneutics and Critical Theory*. Iasington% (acmillan).
- 3arsons, 6. (5//.). Constructivism and Interpretive Theory. In C. (arsh D 2. toler (#ds.), *Theory and Methods in Political Science* (Frd ed., pp. @/09@). Iasington% 3algrave (acmillan).
- 3aterson, 1. E., &horne, . #., 6anam, 6., D 4illings, 6. (5//.). *Meta-Analysis of Qualitative Health Research: A Practical Guide to Meta-Analysis and Meta-Synthesis*. Eondon% age.
- 3atomili, -. , D 7ight, 6. (5//). After Postpositivism: The Promises of Critical Realism. *International Studies Quarterly*, GG(5), 5.F05F>.
- 3ettman, R. (#d.). (.9>9). *Moral Claims in World Affairs*. Eondon% Croom Helm.
- 3hilpott, C. (.99=). On the Gusp of Sovereignty: Lessons from the 15th Century. In E. #. Eugo (#d.), *Sovereignty at the Crossroads: Morality and International Politics in the Post-Cold War Era* (pp. F>0=5). Eanham% Rowman D Littlefield.
- 3hilpott, C. (.999). Westphalia, Authority, and International Society. *Political Studies*, GM, A=0A@9.
- 3hilpott, C. (5//). The Religious Roots of Modern International Relations. *World Politics*, I(5), 5/=05GA.
- 3hilpott, C. (5//.). *Revolutions in Sovereignty: How Ideas Shaped Modern International Relations*. Princeton% Princeton University Press.
- 3lett, -. ., (5//G). *Rhetoric and Renaissance Culture*. Berlin% Walter de Gruyter.
- 3olanyi, C. (.9GG). *The Great Transformation*. Ioston% Beacon Press.
- 3olanyi, C. (.99@). *The Logic of Liberty: Reflections and Rejoinders*. Indianapolis% Liberty Fund.
- 3opper, C. R. (.9FA). *Logik der Forschung*. Vienna% Verlag von Julius Springer.
- 3opper, C. R. (.9=F). *Lectures and Refutations*. Eondon% Routledge and Kegan Paul.
- 3orter, &. (5//.). Can there be Rational Perspectives on International Relations? In R. C. A. Crawford DC. . E. 4arvis (#ds.), *International Relations: Still an American Social Science? Toward Diversity in International Thought* (pp. .F.0 .G9). Albany% SUNY Press.

- 3ouliot, H. (5//G). &he #ssence of 6onstructivism. *3ournal of International Relations and <evelopment*, M(F), F.90FF=.
- 3ouliot, H. (5/.). *International Security in Practice9 The Politics of DJT70Russia <iplomacy*. 6ambridge% 6ambridge *niversity 3ress.
- 3ouliot, H., D Adler, #. (5/..). International 3actices. *International Theory*, A(.), .0F=.
- 3owell, R. (5//F). Beorealism and 2ame &heory. In A. <.-anami (#d.), *Perspectives on Structural Realism* (pp. 9059). 1asingsto!e% 3algrave (acmillan).
- 3owers, 6. -. (5/./). *Ma%ing Sense of Social Theory9 J Practical Introduction* (5nd ed.). 3lymouth% Rowman DEittlefield.
- 3rice, R., D Reus' mit, 6. (.99@). Cangerous Eiaisons8 6ritical International &heory and 6onstructivism. *European 3ournal of International Relations*, G(F), 5A9059G.
- 3sillos, . (5//>). *Philosophy of Science J0*. #dinburgh% #dinburgh *niversity 3ress.
- 3uchala, C. 4. (.99>). ome Bon'7estern 3erspectives on International Relations. *3ournal of Peace Research*, AG(5), .590.FG.
- 3ufendorf, . (.=>A). Ce ystematibus 6ivitatum. <issertationes Jcademicae Selectiones (pp. 5=G0FF/). Eund%-aberegger.
- 3ufendorf, . (>./). *7f the @a/ of Dature and Dations9 Eight Boo%s*. +\$ford% Eichfield.
- 3ufendorf, . (5//5). *The <ivine Feudal @a/9 7r' Iovenants /ith Man%ind' Represented*. (. Qurbuchen, #d.). Indianapolis% Eiberty ,und.
- 3ufendorf, . (5//>). *The Present State of Hermany*. ((. 4. eidler, #d.). Indianapolis% Eiberty ,und.
- Nuine, 7. H. (.9A.). &wo Cogmas of #mpiricism. *The Philosophical Revie/, *)*(.), 5/0GF.
- Nuir!, 4., D Higneswaran, C. (5//A). &he 6onstruction of an #difice% &he tory of a ,irst 2reat Cebate. *Revie/ of International Studies*, A4(/.), @90./>.
- Radnit?!y, 2. (.9>/). *Iontemporary Schools of Metascience* (5nd ed.). A!ademifRrlaget% candinavian *niversity 1oo!s.
- Ramus, 3. (.A>G). *The @ogi%e of the Moste Excellent Philosopher P! Ramus Martyr' De/ly Translated' and in <ivers Places Iorrected' after the Mynde of the Juthor*. Eondon%&h. Hautrolier.
- Ra?, 4. (.99/). *Practical Reason and Dorms*. 3rinceton% 3rinceton *niversity 3ress.
- Rengger, B. (5///a). 3olitical &heory and International Relations% 3romised Eand or #Sit from #den8 *International Jffairs*, M*(G), >AA0>>/. .
- Rengger, B. (5///b). *International Relations' Political Theory and the Pro lem of 7rder9 Beyond International Relations Theory#* Eondon% Routledge.
- Resende' antos, 4. (5//>). *Deorealism' States' and the Modern Mass Jrmy*. 6ambridge% 6ambridge *niversity 3ress.
- Rice, (. A. 4. (5///). 7hat is a cience8 In C. ,. (. trauss D (. lotting (#ds.), *Iontemporary Reflections on the Philosophy of >erman <ooye/eerd* (pp. 5F90 5>.). Eewinston% #dwin (ellen).

- Ricard, J. 4., D Coyle, 4. (5/5). International Relations in Ireland% A survey of Academics. *Irish Political Studies*, (M(5), FA90F>A.
- Riley, 3. (.9>=). &hree >th 6entury 2erman &heorists of ,ederalism% Althusius, -ugo and Eeibni?. *Pu lius*, *(F), >0G..
- Riley, 3. (.9@@). Introduction. In 3. Riley (#d.), @ei ni?9 *Political "ritings* (5nd ed., pp. .0GG). 6ambridge% 6ambridge *niversity 3ress.
- Rit?er, 2. (.9@@). ociological (etatheory% A Cefense of a ubfield by a Celineation of its 3arameters. *Sociological Theory*, *(5), .@>05//.
- Rit?er, 2. (.99/). (etatheori?ing in ociology. *Sociological Forum*, I(.), F0.A.
- Rit?er, 2. (.99.a). Reflections on the Rise of (etatheori?ing in ociology. *Sociological Perspectives*, AG(F), 5F>05G@.
- Rit?er, 2. (.99.b). *Metatheori?ing in Sociology*. &oronto% Ee\$ington 1oo!s.
- Rit?er, 2. (.995a). &he Eegitimation and Institutionalization of (etatheori?ing in ociology. *Sociological Perspectives*, AI(F), AGF0AA/.
- Rit?er, 2. (.995b). (etatheori?ing in ociology% # \$plaining the 6oming of Age. In 2. Rit?er (#d.), *Metatheori?ing* (pp. >05=). Bewbury 3ar!% age.
- Rit?er, 2. (5//.). *Explorations in Social Theory9 From Metatheori?ing to Rationali?ation*. Eondon% age.
- Roosen, 7. (.9@/). #arly (odern Ciplomatic 6eremonial% A ystems Approach. *The 3ournal of Modern >istory*, I((F), GA50A>=.
- Rosenau, 4. B. (5//=). *The Study of "orld Politics' v!49 Theoretical and Methodological Ihallenges*. Bew)or!% Routledge.
- Rosenberg, 4. (.99G). &he International Imagination% IR &theory and L6lassic ocial Analysis.M Millennium9 *Journal of International Studies*, (A(.), @A0./@.
- Roshchin, #. (5/..). (*n)Batural and 6ontractual International ociety% A 6onceptual In"uiry. *European 3ournal of International Relations*, 7nlineFirst, .05F. doi%./...>>:.FAG/=...G55..@
- Rothbard, (. B. (.99@). *The Ethics of@i erty*. Bew)or!% Bew)or! *niversity 3ress.
- Rothbard, (. B. (5//.). *Man' Economy' and State9 JTreatise on Economic Principles* (Gth ed.). Auburn% (ises Institute.
- Rudner, R. . (.9==). *Philosophy of Social Science*. #nglewood 6liffs% 3rentice'-all.
- Ruggie, 4. 2. (.99@). *Ionstructing the "orld Polity9 Essays on International Institutionalization*. Bew)or!% Routledge.
- Zabi[, Q. (5//5). (a!ing a Regional IR &e\$boo!% A Hiew from a 6autious +ptimist. *International Studies Perspectives*, A(F), 5G905AG.
- ampson, A. 1. (5//5). &ropical Anarchy% 7alt?, 7endt, and the 7ay we Imagine International 3olitics. *Jlternatives*, (M(G), G590GA>.
- antoro, (. (5//@). *Id Lias' <iplomacia e <esenvolvimento9 PolTtica Externa Jrgentina de Menem a Qirchner*. l*3#R4.

- aunders, C. (5/.). Hegemon -istory% 3ufendorfPs shifting 3erspectives on ,rance and ,rench 3ower. *"ar' the State and International @a/ in Seventeenth01entury Europe* (pp. 5.505F/). Ashgate.
- chleiermacher, . (.99@). >ermeneutics and Iriticism9 Jnd 7ther "ritings. (A. lowie, #d.). 6ambridge% 6ambridge *niversity 3ress.
- chmidt, 1. 6. (.99G). &he -istoriography of Academic International Relations. *Revie/ of International Studies*, 0(G), FG90F=>.
- chmidt, 1. 6. (.99>). ,urther Ahead or ,urther 1ehind8 &he Cebate over 3ositivism. *Mershon International Studies Revie/*, G4(.), ./>0..5.
- chmidt, 1. 6. (.99@). *The Political <iscourse of Jnarchy9J<isciplinary>istory of International Relations*. Albany% *B) 3ress.
- chmidt, 1. 6. (5//5a). +n the -istory and -istoriography of International Relations. In 7. 6arlsnaes, & Risse, D 1. A. immons (#ds.), >and oo% of *International Relations* (pp. F055). Eondon% age.
- chmidt, 1. 6. (5//5b). &ogether Again% Reuniting 3olitical &heory and International Relations &heory. *British 3ournal of Politics and International Relations*, G(.), ..A0.G/.
- chmidt, 1. 6. (5//=). #pilogue. In <. #. 4Urgensen D & 1. <nudsen (#ds.), *International Relations in Europe9 Traditions' Perspectives and <estinations* (pp. 5AF05=9). Eondon% Routledge.
- chmidt, 1. 6. (#d.). (5/.5). *International Relations and the First Hreat <e ate*. Eondon% Routledge.
- chmidt, . (5/..). &o +rder the (inds of cholars% &he Ciscourse of the 3eace of 7estphalia in International Relations Eiterature. *International Studies Quarterly*, II(F), =/.0=5F.
- chrRder, 3. (.999). &he 6onstitution of the -oly Roman #mpire after .=G@% amuel 3ufendorfPs Assessment in his (on?ambano. *The >istorical 3ournal*, G((G), 9=.0 9@F.
- chweller, R. E. (.99>). Bew Realist Research on Alliances% Refining, Bot Refuting, 7alt?Ps 1alancing 3roposition. *Jmerican Political Science Revie/*, +4(G), 95>0 9F/.
- eidler, (. 4. (5//>). Introduction. In . 3ufendorf, *The Present State of Hermany* (p. i\$0\$\$vii). Indianapolis% Eiberty ,und.
- il, R., D <at?enstein, 3. 4. (5/./). *Beyond Paradigms9 Jnalytic Eclecticism in the Study of "orld Politics*. 1asingsto!e% 3algrave (acmillan.
- il, R., D <at?enstein, 3. 4. (5/..). Ce'6entering, not Ciscarding, the LismsM% ome ,riently Amendments. *International Studies Quarterly*, II(5), G@.0G@A.
- inger, 4. C. (.9=.). &he Eevel'of Analysis 3roblem in International Relations. *"orld Politics*, 4G(.), >>095.
- inger, 4. C. (.9>5a). &he L6orrelates of 7ar 3ro9ectM% Interim Report and Rationale. *"orld Politics*, (G(5), 5GF05>/.

- inger, 4. C. (.9>5b). &heorists and #mpiricists% &he &wo'culture 3problem in International 3olitics. In 4. B. Rosenau, H. Cavies, D (. A. #ast (#ds.), *The Analysis of international politics9 Essays in >onor of >arold and Margaret Sprout* (pp. @/09A). Bew)or!% ,ree 3ress.
- !idmore, 7. (.9>A). *Theoretical Thin%ing in Sociology*. 6ambridge% 6ambridge *niversity 3ress.
- mith, . (.9@>). 3aradigm Cominance in International Relations% &he Cevelopment of International Relations as a ocial cience. *Millennium9 3ournal of International Studies*, 4*(5), .@905/=.
- mith, . (.99A). &he elfImages of a Ciscipline% A 2enalogy of International Relations &heroy. In <. looth D . mith (#ds.), *International Relations Theory Today* (pp. .0F>). 6ambridge% 3olity.
- mith, . (.99=). 3ositivism and leyond. In . mith, <. looth, D (. Qalews!i (#ds.), *International Theory9 Positivism and Beyond* (pp. ..0GG). 6ambridge% 6ambridge *niversity 3ress.
- mith, . (5///). &he Ciscipline of International Relations% till an American ocial cience8 *British 3ournal of Politics and International Relations*, ((F), F>G0G/5.
- mith, . (5//F). Cialogue and the Reinforcement of +rthodo\$y in International Relations. *International Studies Revie/, I(.)*, .5F0.AF.
- mith, . (5/./). Introduction% Civersity and Cisciplinarity in International Relations &heory. In & Cunne, (. <urli, D . mith (#ds.), *International Relations Theories9 <iscipline and <iversity* (5nd ed., pp. .0.F). +\$ford% +\$ford *niversity 3ress.
- ou?a, A. (. (. de Cepresbiteris, E., D (achado, +. & (. (5//G). *J MediaOPo como PrincTprio Educacional9 Bases TeZricas das Jordagens de Reuven Feuerstein*. co 3aulo% enac.
- pegele, R. C. (.9@5). ,rom the Incoherence of ystems &heory to a 3hilosophy of International Relations. *The Revie/ of Politics*, GG(G), AA90A@9.
- pegele, R. C. (.99=). *Political Realism in International Theory*. 6ambridge% 6ambridge *niversity 3ress.
- tafleu, (. C. (.9@5). &heories as Eogically Nualified Artefacts(II). *Philosophia Reformata*, GM, 5/0G/.
- tafleu, (. C. (.9@>). *Theories at "or%9 7n the Structure and Functioning of Theories in Science' in Particular <uring the Iopernican Revolution*. Eanham% *niversity 3ress of America.
- tevenson 4r., 7. R. (5//G). 6alvin and 3olitical Issues. In C. <. (c<im (#d.), *The Iam ridge Iompanion to 3ohn Ialvin* (pp. .>F0.@>). 6ambridge% 6ambridge *niversity 3ress.
- to!er, 2. (5/./). Introduction to 3art 5. In C. (arsh D 2. to!er (#ds.), *Theory and Methods in Political Science* (Frd ed., pp. .@.0.@F). Iasingsto!e% 3algrave (acmillan.
- trauman, 1. (5//@). &he 3eace of 7estphalia as a ecular 6onstitution. *Ionstellations*, 4I(5), .>F0.@@.

- trauss, C. J. (1998). *Reintegrating Social Theory: Reflecting upon Human Society and the Discipline of Sociology*. Lanham: Peter Lang.
- trauss, C. J. (1999). *Philosophy and Discipline of The Disciplines*. Grand Rapids: Eerdmans.
- Wagman, J. (1999). *The Domestic Analogy and World Order Proposals*. Cambridge: Cambridge University Press.
- Wagman, J. (2000). 6. A. 7. (Anning and the Study of International Relations. *Review of International Studies*, 26(1), 90-100.
- Wallerstein, I. (1984). *The Structure of Scientific Theories*. Urbana: University of Illinois Press.
- Wendt, J. (1999). *Ethics, Justice and International Relations: Constructing an International Community*. London: Routledge.
- Wendt, J. (1999). The Contributions of Feminist Theory to International Relations. In J. Smith, D. Roth, D. C. Golewski (Eds.), *International Theory: Positivism and Beyond* (pp. 54-75). Cambridge: Cambridge University Press.
- Wendt, J. (1999). IR Theory after the Cold War. *Review of International Studies*, 25(1), 1-25.
- Wendt, J. (1999). *Philosophy and the Human Sciences: Philosophical Papers*. Cambridge: Cambridge University Press.
- Wendt, J. (1999). Geopolitical Relations in the European Middle Ages: History and Theory. *International Organization*, 53(1), 1-25.
- Wendt, J. (2000). Theorizing the Westphalian System of States: International Relations from Absolutism to Capitalism. *European Journal of International Relations*, 6(1), 43-61.
- Wendt, J. (2000). *The Myth of 'Class' Geopolitics' and the Making of Modern International Relations*. London: Verso.
- Wendt, J. (2001). Geopolitics. *Historical Materialism*, 9(1), 1-25.
- Wendt, J. (2001). Debating the Myth of 'Great Power' Formation, the Interstate System and the Emergence of Capitalism in Europe: A Reconsideration. *International Politics*, 38(1), 1-25.
- Wendt, J. A. (2000). Bringing in Carwin: Revolutionary Theory, Realism, and International Politics. *International Security*, 25(1), 1-25.
- Wendt, J. A. (2001). *War and International Relations: In the Evolutionary Origins of War and Ethnic Conflict*. Easington: University Press of Kentucky.
- Wendt, J. A. (2001). *Fathers of International Thought: The Legacy of Political Theory*. Baton Rouge: Louisiana State University Press.
- Wendt, J. A. (2001). *Thucydides: Peloponnesian War*. London: Duckworth.
- Wendt, J. A. (2001). Seeing IR Differently: Notes from the Third World. *Millennium: Journal of International Studies*, 33(1), 1-25.
- Wendt, J. A. (2001). Hearing Latin American Voices in International Relations Studies. *International Studies Perspectives*, 2(1), 1-25.

- &ic!ner, A. I. (5//@a). Eatin American IR and the 3rimacy of lo pr`ctico. *International Studies Revie/, 4*(G), >FA0>G@.
- &ic!ner, A. I. (5//@b). Eatin American IR and the 3rimacy of Eo 3r`ctico. *International Studies Revie/, 4*(G), >FA0>G@.
- &ic!ner, A. I., D 7aever, +. (#ds.). (5//9). *International Relations Scholarship around the "orld*. Eondon% Routledge.
- &ic!ner, 4. A. (.995). *Hender in International Relations9 Feminist Perspectives on Jchieving Hlo al Security*. Bew)or!% 6olumbia *niversity 3ress.
- &illy, 6. (5//@). *Explaining Social Processes*. 1oulder% 3aradigm.
- &oulmin, . #. (.99/). *1osmopolis9 The >idden Jgenda of Modernity*. Bew)or!% ,ree 3ress.
- &rigg, R. (5//5). *Philosophy Matters9 Jn Introduction to Philosophy*. +\$ford% 1lac!well.
- &ully, 4. (#d.). (.9@9). *Meaning and Iontext9 Quentin S%inner and >is Iritics*. 3rinceton% 3rinceton *niversity 3ress.
- &urton, -. E. (5//.). &he Cearth of Cata in International Relations% A 6ritical Realist olution. (ontrJal% 1I A 7or!ing 2roup on IR as a ocial cience.
- &urton, -. E., D ,reire, E. 2. (5//9). -ybridity under -egemonic Influence in IR cholarship% Realism in outh America. Rio de 4aneiro.
- *to, . (5//A). *1onstructics9 J Methodology of Theory 1onstruction*. Eanham% *niversity 3ress of America.
- Han 1elle, C. A. (5//=). Cinosaur and the Cemocratic 3eace% 3aleontological Eessons for Avoiding the #\$tinction of &heory in 3olitical cience. *International Studies Perspectives, M(F)*, 5@>0F/=.
- Han#vera, . (.99>). *Huide to Methods for Students of Political Science*. Ithaca% 6ornell *niversity 3ress.
- HanCrunen, C. (5//.). *Datural @a/ and the T/o Qingdoms9 J Study in the <velopment of Reformed Social Thought*. 2rand Rapids% #erdmans.
- van ,raassen, I. 6. (.9@/). *The Scientific Image*. +\$ford% +\$ford *niversity 3ress.
- van 2elderen, (. (#d.). (.99F). *The <utch Revolt*. 6ambridge% 6ambridge *niversity 3ress.
- van de -aar, #. (5//9). *1lassical @i eralism and International Relations Theory9 >ume' Smith' Mises' and >aye%*. 1asingsto!e% 3algrave (acmillan).
- Has"ue?, 4. A. (.99A). &he 3ost'3ositivist Cebate% Reconstructing cientific #n"uiry and International Relations &heory after #nlightenmentPs ,all. In <. 1ooth D . mith (#ds.), *International Relations Theory Today* (pp. 5.>05G/). 6ambridge% 3olity.
- Has"ue?, 4. A. (.99>). &he Realist 3aradigm and Cegenerative versus 3rogressive Research 3rograms% An Appraisal of Beotraditional Research on 7alt?Ps 1alancing 3roposition. *Jmerican Political Science Revie/, +4*(G), @9909.5.
- Has"ue?, 4. A. (.999). *The Po/er of Po/er Politics9 From 1lassical Realism to Deotraditionalism*. 6ambridge% 6ambridge *niversity 3ress.

- Higgins, I. (1954). *The British Committee on the Theory of International Relations*. London: Macmillan.
- Higgins, C., & D. Nairn. (1957). The State and the Modern Invention of Intellectual Property as a Critical Theory. *International Relations*, 1(5), 1-10.
- Hinckley, A. (1956). *The Nature of Political Theory*. New York: Oxford University Press.
- Holmes, J. (1957). The Rise and Fall of the Inter-paradigm Debate. In J. Smith, & D. L. Q. Lewis (eds.), *International Theory: Positivism and Beyond* (pp. 1-90). Cambridge: Cambridge University Press.
- Holmes, J. (1957). The Discipline after all these Debates. In J. Smith, & D. L. Q. Lewis (eds.), *International Relations Theories: Discipline and Diversity* (5th ed., pp. 59-106). New York: Oxford University Press.
- Hall, R. L. (1996). *Inside/Outside: International Relations as Political Theory*. Cambridge: Cambridge University Press.
- Hall, R. L. (1998). International Relations and the Concept of the Political. *International Relations Theory Today* (pp. 1-15). Cambridge: Polity.
- Hall, R. L. (1998). The Double Outside of the Modern International. *Ephemera*, 1(1), 1-9.
- Hall, J. (1957). Ruth and Power, Men and Technocrats: Theory and Practice in International Relations. *Review of International Studies*, 1(1), 1-15.
- Hall, J. E. (1995). The Theory, Conceptual Standardization, and the Future of Sociology. In D. R. (ed.), *Metatheorizing* (pp. 1-15). Newbury Park: Sage.
- Hallerstein, I. (1956). *World Systems Analysis: An Introduction*. Durham: Duke University Press.
- Hall, J. (1957). Toward a Science of the Theory. *Integral Review*, 1(1), 1-15.
- Hall, J. (1998). The Progressive Power of Realism. *American Political Science Review*, 92(4), 109-115.
- Hall, J. (1998). International Relations and the World, (any theories. *Foreign Policy*, 44(1) (Special Edition), 190-195.
- Hall, J. (1954). The Relationship between Theory and Policy in International Relations. *Annual Review of Political Science*, 1, 15-30.
- Hall, J., & B. (1998). *Man, the State and the World: A Theoretical Analysis*. New York: Columbia University Press.
- Hall, J., & B. (1998). *Theory of International Politics*. New York: (Crawford Hill).
- Hall, J., & B. (1998). Reflections on Theory of International Politics: A Response to (y Critics. In R. J. (ed.), *Deontology and its Critics*. New York: Columbia University Press.
- Hall, J., & B. (1998). The Origins of War in Realist Theory. *Journal of Interdisciplinary History*, 4(1), 1-15.
- Hall, J., & B. (1998). Realist Thought and Realist Theory. *Journal of International Affairs*, 42(1), 1-15.
- Hall, J., & B. (1998). The Merging Structure of International Politics. *International Security*, 4(5), 1-15.

- Waltz, E. B. (1979). Evaluating theories. *American Political Science Review*, 74(4), 909-929.
- Waltz, E. B. (1978). *Realism and International Politics*. London: Routledge.
- Wallerstein, I. (1984). *Just and Unjust Wars: A Moral Argument with Historical Illustrations*. New York: Basic Books.
- Wallerstein, I. (1984). *Philosophy: The Basics* (6th ed.). London: Routledge.
- Wallerstein, I. (1984). *Conceptual Foundations of Scientific Thought: An Introduction to the Philosophy of Science*. London: Polity Press.
- Waltson, A. (1985). *Diplomacy: The Dialogue between States*. London: Methuen.
- Waltson, A. (1979). Systems of states. *Review of International Studies*, 4*(5), 990-1000.
- Waltson, A. (1995). *The Evolution of International Society: A Comparative and Historical Analysis*. London: Routledge.
- Waltson, A. (1978). Conversations with Hedley. In (Ed.) D. Hatcher, (Eds.), *Remembering Hedley* (pp. 90-100). Canberra: Australian Press.
- Wallerstein, I. (1984). Good Girls, Bitch Girls, and Bad Girls: (a) Paranoia in Robert Schuman's Critique of Feminist International Relations. *Millennium: Journal of International Studies*, 14(5), 607-629.
- Wallerstein, I. (1984). *International Relations Theory: A Critical Introduction* (2nd ed.). London: Routledge.
- Wallerstein, I. (1984). *Sociological Writings* (Ed. by David Held). New York: Continuum.
- Wendt, A. (1985). The Agent-structure problem in International Relations theory. *International Organization*, 39(4), 683-716.
- Wendt, A. (1990). Bridging the theory-gap: a theory gap in International Relations. *Review of International Studies*, 16(4), 507-525.
- Wendt, A. (1995). Anarchy is what states make of it: the social construction of power politics. *International Organization*, 49*(5), 683-716.
- Wendt, A. (1990). An constitution and causation in International Relations. *Review of International Studies*, 16(4), 507-525.
- Wendt, A. (1999). *Social Theory of International Politics*. Cambridge: Cambridge University Press.
- Wendt, A. (1995). Why a world state is inevitable. *European Journal of International Relations*, 1(4), 401-421.
- Wendt, A. (1995). The state as person in International theory. *Review of International Studies*, 21(5), 641-661.
- Wight, M. (1991). Incommensurability and the 'paradigm' of communication: What is the 'redundancy' of Kenneth P. *Millennium: Journal of International Studies*, 21(5), 607-629.
- Wight, M. (1999). The 'root' of the 'Constitution' of the 'Locating Agency' in the Agent-structure problem. *European Journal of International Relations*, 5(1), 101-115.

- Tight, G. (5///). Interpretation All the 7ay Cown8 A Reply to Ro\$anne Eynn Coty. *European Journal of International Relations*, *(F), G5F0GF/.
- Tight, G. (5//5). 3hilosophy of ocial cience and International Relations. In 7. 6arlsnaes, &. Risse, D 1. A. immons (#ds.), >and oo% of *International Relations* (pp. 5F0A5). Eondon% age.
- Tight, G. (5//G). &heori?ing the (echanisms of 6onceptual and emiotic pace. *Philosophy of the Social Sciences*, AG(5), 5@F0599.
- Tight, G. (5//=). *Jgents' Structures and International Relations9 Politics as 7ntology*. 6ambridge% 6ambridge *niversity 3ress.
- Tight, G. (5//>). A Response to ,riedrich <ratochwil% 7hy hooting the (essenger does not (a!e the 1ad Bews 2o AwayS *Journal of International Relations and <velopment*, 4)(F), F/.0F.A.
- Tight, G. (5//9). &he Impossible Cream% &heorising in International Relations without the 3hilosophy of cience. *ISJ D)th Jnnual Ionvention* (p. F9). Bew)or!.
- Tight, (. (.9G=). *Po/er Politics* (.st ed.). Eondon% Royal Institute of International Affairs.
- Tight, (. (.9==). 7hy is there no International &heory8 In -. 1utterfield D (. Tight (#ds.), <iplomatic Investigations9 *Essays in the Theory of International Politics* (pp. .>0FG). Eondon% Allen D *nwin.
- Tight, (. (.9>>). *Systems of States*. (-. 1ull, #d.). Eeicester% Eeicester *niversity 3ress.
- Tight, (. (.9@>). An Anatomy of International &hought. *Revie/ of International Studies*, 4A(F), 55.055>
- Tight, (. (.99.). *International Theory9 The Three Traditions*. (2. Tight D 1. 3orter, #ds.). Eeicester% Eeicester *niversity 3ress.
- Tight, (. (5//A). *Four Seminal Thin%ers in International Theory9 Machiavelli' Hrotius' Qant' and Ma??ini*. (2. Tight D 1. 3orter, #ds.). +\$ford% +\$ford *niversity 3ress.
- Williams, -. (.995). *International Relations in Political Theory*. (ilton <eynes% +pen *niversity 3ress.
- Wilson, 3. (5///). 6arr and -is #arly 6ritics% Responses to L&he &wenty yearsM 6risisP, .9F9'G=. In (. 6o\$ (#d.), E!>! *Iarr9 J Iritical Jppraisal* (pp. . =A0.9>). Eondon% 3algrave.
- Wilson, 3. -. (.999). *The >oly Roman Empire9 4G+I04,)*. 1asingsto!e% (acmillan.*
- Wilson, 3. -. (5//@). *Europe's Tragedy9 a >istory of the Thirty Eears "ar*. Eondon% Allen Eane.
- Winch, 3. (.9A@). *The Idea of a Social Science and its Relation to Philosophy*. Eondon% Routledge and <egan 3aul.
- Titte 4r., 4. (5//>). *The Reformation of Rights9 @a' Religion' and >uman Rights in Early Modern lalvinism*. 6ambridge% 6ambridge *niversity 3ress.
- Wohlforth, 7. 6. (.999). &he tability of a *nipolar 7orld. *International Security*, G(5), A0G..

- Wohlfarth, W. G., Eittle, R., Kaufman, J. A., Lang, C., Jones, G. A., -ui, H. & 'l.,
#c!stein, A., et al. (5//>). &esting balance'of power &heory in World -istory.
European Journal of International Relations, 4A(5), .AA0.@A.
- Woods, B. (#d.). (.99=). *Explaining International Relations since 4+GI.* +\$ford%
+\$ford *niversity 3ress.
- Qahar, #. (.9>F). Why did #insteinPs 3rogramme upersede Eorent?'s8 (I). *The British
Journal for the Philosophy of Science*, (G(5), 9A0.5F.
- Qalews!i, (. (.99=). LAll these &heories yet the lodies <eep 3iling up M% &heories,
&heorists, &heorising. In . mith, <. looth, D (. Qalews!i (#ds.), *International
Theory9 Positivism and Beyond* (pp. FG/0FAF). 6ambridge% 6ambridge *niversity
3ress.
- Qara!ol, A. (5/..). &heori?ing International Relations% 3olitics vs. 3hilosophy.
International Studies Revie/, 4A(G), =G>0=AF.
- Qehfuss, (. (5//5). *Constructivism in International Relations9 The Politics of Identity*.
6ambridge% 6ambridge *niversity 3ress.
- Qhao, . (.99.). (etatheory, (etamethod, (eta'Cata'Analysis% 7hat, 7hy, and -ow8
Sociological Perspectives, AG(F), F>>0F9/.
- Qhao, . (5//G). (etatheory. *Encyclopedia of Social Theory*. age.
- Qurbuchen, . (.99@). amuel 3ufendorf and the ,oundation of (odern Batural Eaw%
An Account of the tate of Research and #ditions. *entral European >istory*,
A4(G), G.F0G5@.
- Qurbuchen, . (5//5). Introduction. In . 3ufendorf, *The <ivine Feudal @a/9 7r'
Iovenants /ith Man%ind' Represented*. Indianapolis% Eiberty ,und.

Segbers/
Dyllick-Brenzinger/Hoffmann/Mauersberger

**Global Politics:
How to Use and Apply Theories of
International Relations**

56_{/2006}

**ARBEITSPAPIERE
DES OSTEUROPA-INSTITUTS DER FREIEN UNIVERSITÄT BERLIN
ARBEITSSCHWERPUNKT POLITIK**

Mit der Reihe „**Arbeitspapiere des Osteuropa-Instituts**“ stellt der Arbeitsschwerpunkt Politik aktuelle Ergebnisse aus seiner Arbeit der Öffentlichkeit zur Verfügung.

Die Arbeitspapiere erscheinen in unregelmäßiger Folge. Einzelne Papiere können gegen eine Schutzgebühr bezogen werden bei:

Osteuropa-Institut der Freien Universität Berlin
Garystr. 55
14195 Berlin
<http://pol.oei.fu-berlin.de/>

Tel.: ++49 30 838 54058
52075
Fax: ++49 30 838 53616

Arbeitspapiere des Osteuropa-Instituts
der Freien Universität Berlin
Arbeitsbereich Politik und Gesellschaft

**Segbers/
Dyllick-Brenzinger/Hoffmann/Mauersberger**

Global Politics: How to Use and Apply Theories of International Relations

Heft 56/2006

© 2006 by Klaus Segbers, Peter Dyllick-Brenzinger,
Katharina Hoffmann, Christof Mauersberger

Osteuropa-Institut der Freien Universität Berlin
Arbeitsbereich Politik und Gesellschaft
Herausgeber: Klaus Segbers

KLAUS SEGBERS:

INTRODUCTION	3
---------------------------	----------

KLAUS SEGBERS:

THE GAME OF THORIES	5
POSITIVISM – CONSTRUCTIVISM/ META THEORIES	5
REALISM/ NEO-REALISM	13
INSTITUTIONALISM	15
THIRD (FIFTH) IMAGE APPROACHES	18
THE LIBERAL/ PLURALISTIC AND DOMESTIC STRUCTURE APPROACHES	21
COGNITIVE APPROACHES	24

PETER DYLLICK-BREZINGER AND CHRISTOF MAUERSBERGER:

ISRAELI-PALESTINIAN CONFLICT	27
AN INTRODUCTION TO THE ISRAELI-PALASTINIAN CONFLICT, FOLLOWING WALTZ' LEVELS OF ANALYSIS	27
INSTITUTIONALISM IN THE MIDDLE EAST CONFLICT	31
FROM DOVE TO HAWK – CHANGES EXPLAINED	33
BARAK AND ARAFAT PLAYING (TWO-LEVEL-)GAMES	36
COGNITIVE APPROACHES IN THE MIDDLE EAST CONFLICT	39

KATHARINA HOFFMANN:

IRAQ CONFLICT	43
CONFLICT SETTING: IRAQ – AN INTRODUCTION	43
DO IDEAS MATTER?	47
DECISION OVER A “SECOND RESOLUTION” - A PRISONER’S DILEMMA?	50

CHRISTINE SCHUSTER:

THE CONFLICT IN THE DEMOCRATIC REPUBLIC OF CONGO	55
“THE CONFLICT OF THE GREAT LAKES”	55
IS ANYBODY STILL A STATE IN CONGO?	60
APPLICATION OF PUTNAM’S “TWO-LEVEL-GAME” THEORY TO THE CONFLICT IN DRC	63

ONDŘEJ SPAČEK:

THE CRISIS IN COLUMBIA	68
COLUMBIA – AN INTRODUCTION	68
CONSTRUCTIVISM – THE ALL EXPLAINING WONDER THEORY?	71
GLOBALIZATION	73

INTRODUCTION

BY KLAUS SEGBERS

This working paper presents the results of my introductory lecture into global politics in the summer term of 2004 at Freie Universität Berlin. The organizing idea is to provide future students with some basic knowledge and tools for being prepared for future, similar lectures.

The idea of this lecture is to give the students an overview of approaches and issues relevant for International Relations (IR). This is implemented in the following way. During the first hour, I present and explain relevant approaches to and issues of IR. Afterwards, for the second hour, the students disperse into groups and work on given questions related to the main topic, i. e. mostly to the theory discussed in that session. The respective theory has to be applied to a case, i.e. to a country or conflict. The countries / conflicts are stable within one group over the whole term. Finally, in the third hour, reports from the working groups are presented to the plenary session.

This working paper aims at both – presenting the theories, and giving an idea of the work in the working groups. The following pages will offer brief sketches of the most relevant approaches to interpreting international relations / global politics. We start with looking into the two macro-approaches of social sciences, namely positivism and constructivism. Then we proceed to neo/ realism, institutionalism, world systems, liberalism / pluralism, and cognitivism.

Afterwards, we present examples of the tasks, addressed by the students of my lecture in the summer term 2004. They had to apply those theories to their respective region and conflicts.

The structure of the lectures differs somewhat from this outline and will be presented briefly. I start with the two meta-theories: positivism and constructivism. The IR theories are arranged re. their views on and from different levels of analysis. First come the global (or world) system theories (5th or system level perspective), followed

by institutionalism and regime theories (explaining 3rd (state) - and 4th (supra-state) level phenomena from a 3rd (state) image perspective). Next in line are realism and neo-realism, addressing 3rd level behavior from either a 1st (individual) or 5th level outlook. Then we have theories like two-level games and second image reversed explaining 3rd level outcomes from a 2nd (domestic) level. This is followed by pluralist, liberal and domestic structure approaches explaining state actions (3rd level) from a 2nd level of analysis. Finally, there are cognitivist theories looking into 1st image factors such as mind sets for explaining IR.

THE GAME OF THORIES

BY KLAUS SEGBERS

Substantial Theories of IR	Meta-Theories		
	Positivism / Ratchoice		Reflectivism / Constructivism
	World Systems	Institutions	
	Cognitive Approaches		
	(Neo-) Realism	Liberalism / domestic approaches	

POSITIVISM – CONSTRUCTIVISM/ META THEORIES

Before we come to briefly presenting theories interpreting global politics, we have to make an important digression. We have to think about what we are doing here, and how we are doing it – making sense of global politics. This is related to very general and important questions related to **social sciences**. Two things have to be made clear. First, it is not possible to think about anything (at least in scientific terms) without making some very **basic assumptions** about how this will be organized. So the apparently attractive option sometimes heard in seminars “let’s do this without theories” is not workable. Even when these basic assumptions are not explicated, they are still there. We can’t do without them. And so it is better to make them direct and open.

Second, **theories** are not something gruesome. There is no horror, no unsurmountable threshold involved. They are not really difficult to understand once you accept them as necessary and you start trying. And in science, we can’t do without them. Theories are not something scary but simply general assumptions about how factor a causes factor b. That is what science is about: **explaining things**. That sounds natural, but it is not. Many students assume that describing things is science, too. But even when we try to describe something we again cannot avoid using a guiding assumption as a roadmap (what data to look at, what cases to

compare etc.). And secondly, most phenomena in social / political sciences are already available as data, and there is no need to re-collect them. But when they are not available, the endeavor to do that is way too big for a student's thesis.

The following pages are on how to use theories, and on what the most relevant theories are. For this purpose, it is useful to introduce a concept which helps to sort out different theories. This is the concept of **levels of analysis** (loa), which means: on what analytical level are we working when analyzing a given problem? It is obvious that when we are moving quickly back and forth from one level to another, we finally do not really know which is the most relevant one. So we first have to define a loa, and then stick to it until we formally may want to change it to another loa.

There are different suggestions re. which, how many loas are required. This is what I propose (and will stick to):

- 1st level:** individual actor/ agent / decision maker
- 2nd level:** societal / social/ group/ transnational/ subnational actors ("domestic" sphere)
- 3rd level:** unit / state/ government
- 4th level:** supranational entities
- 5th level:** global/ international system.

These are 2 more levels than suggested by Kenneth Waltz in his influential introduction on "images" (levels) of IR analysis¹. But I maintain that this is useful and legitimate. By the way: you see that we can be creative with theories. We can develop them. We may touch and modify them.

Now before we move on to the theories, there is another concept that requires explanation (and then will be used in combination with the loas): **variables**. "Variables" is a term for groups of factors in "real life" that cause other factors, or are being caused by other factors. There are two basic types of variables: **dependent variables** (DV) – those factors being caused – and **independent variables** – the

causing factors (IV's). There are more groups of factors in the literature on methods, but these two core groups will suffice here.

Obviously, variables must be able to vary; if not, they cannot serve as variables. So we have to check in every instance when we define factors as variables if they have the potential to vary. Also, variables rarely can be observed directly. So we have to look at them on a somewhat more concrete level, namely that of **indicators** or proxies. There are many relevant indicators for IR questions for which we do have very good data, often in data banks and on the internet (demographic data, the Human Development Index, data on FDI, on conflicts and wars, on migration flows etc.). A good variable that cannot be operationalized with indicators is almost worthless.

Once we have defined variables and know how to measure them, we are able to formulate a **hypothesis**; this is a statement re. how a given independent variable shapes the defined dependent variable. There should be as many hypotheses as IV's². Basically, a hypothesis is a less generalized, i.e. more concrete theory; or a theory applied to a precise case.

Each of the IR theories to be described below can be classified by looking at its assumptions re. on which the global politics phenomena take place, and from which they should be best explained.

But before we address the so-called IR theories directly there is another step we should make, and this is on a **meta-level of reasoning**. It is not "clear" automatically how we are thinking, or how we should do that. Actually, there is a massive amount of literature exploring this field of how we "think", and on how we interact with "reality". These two terms are not self-explaining. Neurobiologists, evolutionary biologists, linguists, philosophers and other people in other disciplines have a fascinating, but also a hard time making "sense" of our way of "making sense".

¹ Waltz, Kenneth: Man, the State and War, New York 1959.

² For more details on variables and theories, see van Evera, Stephen: Guide to Methods for Students of Political Science, Ithaca NY 1997, p. 7-49.

There are two basic ways of reasoning, and accordingly two main groups of researchers. This constitutes the important divide between the **two big camps** – positivists and constructivists. The decisive difference between them is how their representatives look at “reality”. While positivists maintain that reality does exist “outside” of their work (is exogenous), and that one can catch it by designing research procedures in a proper way (or at least that it makes sense to assume that we can do that), constructivists deny these assumptions or at least say that there is no evidence to assume the accessibility of reality. While this divide may seem quite academic – we all make reality assumptions or we could not act daily in our respective contexts – this academic debate on the nature of reality is very important indeed. And, most probably, it cannot be resolved in a convincing intersubjective way. So at the end of the day this is one point where you – we all – have to take a normative decision, i.e. a decision that in itself cannot be legitimized by science, but it rather is derived from plausibility, norms, and values.

Positivists take the position that there are actors which have preferences and interests and that they act accordingly to achieve them. All these elements – actors, preferences and interests – can be described and analyzed. Also, how these actors act and shape events, and processes, can be tracked. There is an inbuilt tension between actors / agents and structures / institutions, but basically it is an interaction that can be handled productively.

Constructivists deny most or all of this. They say that all preferences and interests are, or may be, fluent and shifting, and that they are all subject to changes during the process of social interaction and also of research. Therefore, one cannot take them for granted and certainly not as a starting point and a stable orientation for research.

When we look at these metatheories from the position of loas, we can see that loas in themselves are basically an invention by positivists. Positivists can act from all of the above mentioned loas. Constructivists either would avoid framing research questions in loas, or they would say that their main loa is intersubjectivity – no matter where it occurs.

For this working paper on IR theories and their application, you will detect that there is NO principal decision on which of these two ways of thinking, and doing IR is “better”. This is something you have to decide by yourself. But there is a strong tendency, at least from my side, to accept **positivist assumptions** pragmatically for practical purposes. Using them will allow you to build meaningful and practical research designs (study proposals). Using them also will allow you to be prepared for most of the practical contexts of your professional positions you later will apply for, and work in.

The core of many positivist, actor-oriented approaches is a rationality assumption. The methodical principle behind it is called **rational choice**. That means that actors do act rationally NOT in the sense of acting reasonably or “good”, but they act in a goal-oriented manner. The actions are input-rational, not necessarily out-put reasonable. The actors have the same pattern of doing things – by the way, in all spheres of life. They have fixed and mostly stable preferences (like survival, reproduction, resource accumulation, security/ protection, maybe cognitive consistency, see below) and act in a given situation to achieve their preferences. So they are utility maximizers (and still can be risk-prone or risk-averse). They screen the information at hand and then arrange the available options hierarchically, and act accordingly. While there certainly are limitations of information, of processing, of resource availability and of other constraints (like rules / institutions), the basic modus operandi is goal-oriented (a statement constructivists would attack).

This principle is supposedly universal and across cultures, disciplines and activities. The economic success of a product, the political success of a political candidate and the international success of a state are all dependent on competition and on selections made by many individual actors: by consumers on the market, voters in elections, and by other states’ behavior in conflicts. While the outcome may (and often will) differ from the intentions of most or even all those actors, the **organizational principle** of how they make their decisions, and that their preferences are their roadmap for action is relevant for analyzing and even predicting politics (and other social activities).

Positivists and rat-choice proponents are supposedly particularly effective in addressing decision theories (acting under / within constraints of a given situation), crisis behavior (Cuba, Berlin, 9/11 crises etc.), games (strategic action under constraints imposed by other actors / interdependent choices) and when taking up collective action problems. A sub-field of positivism / rat-choice is game theories.

But this is not a statement on the eternal merits of positivism. We – and I – just don't know. Constructivists are very strong in some important regards. First of all, they flatly make clear the inbuilt and quite often (though not here) hidden assumptions of positivist thinking. They can indicate and delegitimize the power potential of dominant positivist discourses. They can show that, and why, researchers are not neutral but active participants of discourse construction. They can demonstrate the crucial role of communication and texts for social sciences, politics, and IR. So this is definitely a very relevant method of reasoning (and questioning hegemonic ways of reasoning). They focus on topics like the role of social factors, norms and practices, and the importance of cultural aspects, habits and traditions.

Post-positivists put forward relevant questions to positivists and rat-choice proponents: Do actors have sufficient information to make their choices? What exactly does „rational“ mean? Is there a universal definition of „utility“? Is the assumption of exogenously set preferences correct? How about acting in situations of uncertainty? What about shifting (unstable) interests / priorities, especially under the influence of communication? Are there no non-hierarchical, cooperative situations? Can they be achieved by communication?

Constructivists also have an impressive record of thinking behind – or underneath – apparent “reason” and the dominant narrative of modernity. They came up with the first skeptical questions toward modernity after World War I, the next principal questions after World War II (“Dialectics of Enlightenment”), they used the power-deconstructing writings of Foucault, the text-deconstructions of Derrida and Baudrillard, and the discussions around risk-societies and new risks in these days of the early 21st century. Since we all are in a situation where modernity is attacked on different fronts – from religious fundamentalists of all sorts, from disoriented and

dislocated and discouraged social groups, and is discredited by ad-hocist policies, constructivists have to be aware of the potentials of their de-legitimizing potentials of their legitimate discourse critique.

There are also tough questions for constructivists. Is there anything beyond the „subject“? Is there more than to ramble and to waffle? Are there any rules of discourse, any rule enforcement, anything beyond „anything goes“? What exactly can we do to explain ir / global politics applying these critical tools?

To sum up, meta-theories propose more general assumptions about
... the world and the nature of objects (ontology),
... how we know that and what we know about this reality (epistemology), and
... how to get (to) this knowledge (methodology).

While positivists stand in the tradition of enlightenment, modernity and rationality, constructivists mostly would not deny the merits of these projects (or narratives) but rather focus on their limits, and on the dangers of exclusiveness.

A brief dictionary for understanding meta-theories:

Ontology: **A branch of science theory concerned with the nature and relations of being, or things which exist.**

Epistemology: **A branch of science theory concerned with knowledge or ways of knowing, particularly in the context of the limits or validity of the various ways of knowing.**

Hermeneutics: **The science of interpretation and explanation; exegesis.**

Heuristics: **Strategies of how to search for solutions.**

Historically, there were some so-called great debates in IR development. These debates served, and serve, as organizing principles for the relevant epistemological conflicts being addressed in and between university departments, research journals and, to some extent, the media.

These debates, as they are mostly portrayed, were those between

-- idealists / institutionalists and realists (since World War I),

-- behaviorists / empiricists and hermeneutical approaches (since the 60s),

-- and then between all positivist and the others (de/constructivist, reflectivist, gender-oriented) positions, creating “linguistic” and “cultural” turns.

Now we will present five groups or IR theories that do generally follow positivist rules.

They are mainly interested in the following **questions**:

- How are political decisions generated by actors?
- How do we explain political outcomes produced by these decisions?
- Under what circumstances is cooperation possible, and when and why does it fail?
- How do rules / institutions / regimes work?

REALISM/ NEO-REALISM

There is no one single coherent definition of “realism.” There are at least two main brands: classical realism, and neo-realism. Classical realism was not very precise in defining its core assumptions. But a certain understanding of “man” –human beings – , human nature, and of their eternal aspirations for power played a crucial role. This proposition looked increasingly odd, and in the 60s and 70s classical realism was sidelined. Apparently, more recent findings of evolutionary biology may lead to a renaissance of this concept, in combination with cognitive approaches (see below).

Some 25 years ago, classical realism was mostly replaced by neo-realism. The original thinker behind neo-realism was Kenneth Waltz. Neo-realism was presented as an alternative to all other IR theories because it rested on “real” scientific assumptions. While this claim was and is questionable, neo-realism became quite influential. It rests on the assumption that states behave in the framework of the international system in the same way as producers and consumers do in economic markets. The “organizing” principle in both cases is, according to Waltz, anarchy, and the players behave accordingly. First of all they develop survival strategies.

For a long time, international politics was defined and understood as politics between states. Accordingly, the discipline of IR required the existence and interaction of and between states. Consequently, reflections on the state are quite old: theories of the state, state-based law, etc. But we should note that clear definitions of “the” state always were not so easy to construct.

The modern state was the core element of the so-called Westphalian system (1648-1991), but it was preceded by the ancient Greek polis and other state-like entities. Attributes of states were the existence of an central administration, territoriality, internal sovereignty (esp. power monopoly), social homogenization, external sovereignty, citizenship/ Staatsbürgerschaft, national identity, borders (indicating distinct domestic/ foreign spheres), and specific state symbols.

Loas: Realists were not quite specific here, but mostly they used the 1st level to explain outcomes on the 3rd level (2nd in Waltz’s terminology). Neorealists focus on

the 5th (in their parlance - 3rd) level, especially on anarchy, to explain system-level behavior (states, i.e. 3rd level in our terminology).

Core ideas of realism are that the environment of states is anarchic and, in a way, dangerous. States are understood as unitary actors. Governments striving for survival had to follow the principle of self-help. This makes cooperation neither likely nor stable and, therefore, not reasonable. Another central realist assumption is that states are only interested in relative gains (compared to other rival states), not in absolute advantages.

Core thinkers of classical realism were Thucydides, Machiavelli and Hobbes. They were followed in the 20th century by Morgenthau, Schelling and others. Neo-realism is represented by authors like Waltz and Mearsheimer.

A decisive difference between them is that classical realism rested mostly on assumptions re. the 1st level, and how that affected the 2nd (state) level, whereas neo-realism focuses on the 3rd level to explain the behavior of states on the 2nd level.

Critique: Today, in global times, it is more problematic than ever to maintain that states are the only, or even the main players on the globe. Also, it seems that realists were not particularly successful in predicting – or even explaining - the end of the Cold War.

Also, the unitary actor thesis is hardly “realistic”. Still, there are some interesting contributions from neo-realist thinkers. The prediction in the early 1990s that the bipolar Cold War system was more stable than the new (dis) order that replaced it has some merit.

Tasks for work grouping:

1. Try to model your region / conflict first with a realist and then with a neo-realist approach.
2. Why does anyone still engage in nation building if states are eroding?
3. Are realists and neorealists normative?

INSTITUTIONALISM

There is some confusion re. how to label this approach. Suggestions are: institutionalism, functionalism, idealism... Others talk about „neoliberal institutionalism. “ For our purpose here, I call this approach institutionalism. It is very close to regime theories.

Definition: What is institutionalism about? This approach wants to explain cooperation between (state) actors and coordination of their actions. It also addresses the core question of in/stability and in/security. Institutionalists maintain that cooperation between states / governments is possible and may be effective.

Loa: Institutionalists try to explain 3rd level performances from a 3rd or 4th level perspective.

Core ideas: It probably makes sense to start with a look at institutions in general. What are they? Institutions can be best understood as rules of the game. They may be formal or informal, legal or illegal, effective or not. Rules labeled as institutions are constraints – and opportunities – for actors. They should not be confused with organizations.

Institutions are formal and informal rules that constrain individual behavior and shape human interaction (Douglass North).

The concept of institutions was developed by economists and constituted the approach of neo-institutionalism. Increasingly, it is applied by political scientists as well. Related terms and concepts are transaction costs, principal-agent relations, institutional change, path dependency and regimes.

Regimes are the institutions of IR. They are sets of implicit or explicit principles, norms, rules, and decision making procedures around which actors' expectations converge in a given area of international relations (Stephen Krasner).

Basic assumptions of institutionalist thinkers for international relations are the following. Institutionalists have a similar understanding of the international system (anarchic) and the state (unitary) as realists. But they are saying that the state can be embedded in rules and act in such a way that its inherent behavior – utility maximizing strategies leading to permanent instability for everyone (security dilemma) – can be overcome by utility maximizing – leading to increasing interdependence b/w states, thereby producing stability and cooperation. A compatible way of reasoning is that states accept rules / regimes even in those cases when their short-term effects are detrimental for them because (only) this guarantees that other participants also will accept these rules in the future. So the potential negative future impact of defection today - the so-called shadow of the future – ensures rule acceptance now.

Institutionalists realize that modern societies are characterized by complex schemes of division of labor. This makes them vulnerable for dysfunctions, for example by attacks. So governments may develop an interest in intertwining and integrating some of their functional spheres with other societies – resulting in networks, and in increasing mutual vulnerability.

These kind of functional approaches are not necessarily idealistic (an early 20th century term for the assumptions that wars can be avoided by international institutions). It is especially worthwhile to note that we do not need to call for altruistic principles, but rather expect common interests because there are common institutions and regimes leading to reduced TACs and making defections difficult.

So regimes and institutions can be useful, according to their proponents, because they do not rest on appeals and values, but because they rather serve mutual interests. Therefore, they can be integrated into interests, or even preferences, of state actors – who expect utilities from regimes.

Basic thinkers/ proponents: The economist and Nobel laureate Douglass North is the most influential thinker on institutions. But there is a whole school of neo-institutionalists. Stephen Krasner coined the leading definition for IR regimes.

Critique: Institutionalists have problems in explaining the breakdown of institutions, or their ineffectiveness. When observing rules is such a convincing concept, why is there not more rule-observance in global politics?

Tasks for work grouping:

1. What relevant institutions can you detect in your region? In what policy fields?
2. Do international laws and regimes constrain (or enable) the behavior of actors? How?

THIRD (FIFTH) IMAGE APPROACHES

There are at least nine different approaches explaining processes and events on levels 1 to 3/5 (1 to x) with third image causation. There are more or less Marxist theories, like Imperialism (Lenin), social hegemony (Gramsci), world systems theories (Immanuel Wallerstein), and the critical theory of world orders (Robert Cox).

Then we do have non-Marxist approaches which belong into the 5th level group, namely international Political Economy (IPE, Susan Strange et al.), neo-realism (Kenneth Waltz and others), some globalization theories, some global culture theories (Samuel Huntington), and some game theories that can be applied to IR.

Loa: All these approaches look from a 5th level (formerly 3rd level) to explain outcomes on other levels.

These approaches differ significantly in their relevance today. Not all of them can be explained here. Globalization is one of the currently most important fields of analysis, IPE as well. Neo-Realism was addressed in connection with Realism. Civilization has to be dealt with in the culture session. Imperialism is „out. “

Gramscianism is not of much relevance today in IR.

Core ideas: There are some central ideas more or less common to most 3rd level approaches. Obviously, the relevant level of analysis is the 3rd level. For most of these theories, the world system is primarily characterized by global capitalism. The world system is, so to say, the playground of world capitalism, and is being shaped and structured by world capitalism. Because of the inherent dynamics of wc, the ws is „injust“. It is heterogeneous, and the decisive marker is the capability of actors / states to belong to the „core. “

There is a social dimension of interaction. Relations between social groups are as – or even more – important, as diplomatic relations between states and governments.

One important theoretical source of many 3rd level approaches is Marxism. This started with the critique of imperialism and was further developed into Leninism. These theoretical currents were more / or less popular (at different times?) depending on crisis situations. Antonio Gramsci's concept of hegemony was a quite sophisticated attempt to adapt Marxism to the conditions in the middle decades of the 20th century. He included the concept of culture into his thinking.

Another influential theoretical background came from history, particularly from the French group around the journal "Annales". The historical Annales school was focusing not on individuals, emperors, or single events, but suggesting a multiple time model: the *longue duree*, medium term time frames (*conjunctures*) and events.

Core thinkers: These are primarily Immanuel Wallerstein and Robert Cox. Wallerstein is probably the most influential thinker in the group of 3rd level approaches to IR. He developed a concept called the world system approach. This is an all-encompassing idea operating on the level of world empires and world economies. The world system we are living in is determined by the global economy, namely: the capitalist economy. This global capitalist economy produces three different, but also interrelated spaces: cores, peripheries, and semi-peripheries.

Core members are characterized by democratic governments, relatively high wages, the import of resources, the export of manufactured goods, and efficient welfare services. Periphery members are characterized by the absence of these qualities. Semi-periphery regions are an intermediate space between the two others. A central question here is, of course, if and how a political entity can "migrate" from one group to another.

In addition to this spatial organization, there is also a temporal dimension of world systems such as cycles, trends, waves, crises etc. The role of states in this system is important.

Robert Cox is another good example for interesting 3rd level approaches.

He is quite reflective on theorizing: He discriminates between problem-solving theories, and critical theories.

The basic analytical category for Cox is historical structures. This term is defined as a „framework for action“, as a „particular configuration of forces“. In any given historical structure, writes Cox, there are three categories of forces which interact: Ideas, institutions, and material capabilities³. These social structures are „located“ or operative on („applied to“) three different levels: Social forces, states and world orders⁴. This concept developed by Cox is quite productive. He demonstrates what can be done with 3rd level thinking.

The most dominant approaches today are related to globalization as a process. The reduction of the meaning of borders, the changes on the regulatory capabilities of governments and states, the increasing importance of flows and scapes, ongoing ways of compression of time and space, digitalization and many other variables and phenomena have certainly a major impact on global politics. There are too many authors and writings on this subject to enumerate them here. For more details and analysis, I suggest to turn to the papers and books written by Appadurai, Castells, McGrew and Scholte, among others.

Critique: Wallerstein has problems predicting if, and under what conditions states may move up-or downwards in his hierarchy. Cox shows implicitly that “adequate” theories are exposed to the danger of “overcomplexity.”

Tasks for work grouping:

Please apply Wallerstein, Cox or Game theory to your region/ conflict.

1. What is “core,” what is – “periphery?” How can a unit move upward / downward?
2. What are the historical structures that frame the playground for the relevant actors?

³ Cox, Robert W.: Social Forces, States and World Orders. Beyond the International Relations Theory, in: Keohane, Robert O.: Neorealism and its Critics, New York: Columbia University Press 1986, p.218.

⁴ *ibid.*, p.221.

THE LIBERAL/ PLURALISTIC AND DOMESTIC STRUCTURE APPROACHES

The analytical focus of liberalism / pluralism shifts toward domestic structures and actors: What is their impact on state behavior? What is their own transnational role?

In this case, the dependent variables are, still, aspects of international or world politics. But the domestic context (structures) will be brought in as independent variables.

Loas: All these approaches start at the 2nd (domestic / societal) level to explain causal effects on the 3rd (state) and 4th levels.

This is different from explaining domestic events / processes with international factors (like globalization). And this is different, too, from explaining state behavior by system-level qualities (anarchy; distribution of capabilities / information).

Liberalism

Liberalism is not just one concept. It has broad implications and many different meanings. Liberalism in economics is different from political liberalism in the U.S. also, liberal parties are quite different (see Germany, Austria and Russia).

In an IR context, liberalism means that democratic states don't fight each other – this is the theory of democratic peace. But note: democracies do fight non-democracies.

The basic tenet of liberal approaches in IR is that „state-society relations - the relationship of states to the domestic and transnational social context in which they are embedded - have a fundamental impact on state behavior in world politics“ (Andrew Moravcsik).

There are important implications to draw from this kind of reasoning. This approach is not compatible with states as „black boxes,“ states as effective resource mobilizers, states as “containers,” states as unitary actors, states as rational decision makers (though it may be compatible with rat. choice).

Other implications are that societies and social actors are the relevant level / object of analysis. They act and interact, they build coalitions, they lobby, they put pressure on bureaucracies, they act – collectively or individually.

This is what being „liberal“ or „plural“ means in this IR context.

Core assumptions: There is a primacy of societal actors, of individuals and social groups. Political actors are dependent on election cycles (and, in general, on time). The proper definition of their interests is central.

„States ... are not (unitary) actor(s) but representative institution(s) constantly subject to capture and recapture, construction and reconstruction by coalitions of social actors.“ (Moravcsik 1997, 518)

Ideational liberalism

This approach is about the role of ideas and ideologies. It looks at societal preferences concerning the scope and content of a „nation.“ It is interested in the commitment of individuals and social groups to particular ideas and political institutions (regime types).

Commercial liberalism

This approach is about the relevance of patterns of market incentives for domestic and transnational economic actors. The matter here is not only free trade. The underlying thesis is more general: The greater the economic benefits for private actors, the greater their incentive to press governments to facilitate such transactions.

Republican liberalism

This interpretation addresses the mode of domestic political representation which determines whose social preferences are institutionally privileged. Rent-seeking is an important mode of action in this framework.

Two-level games

Another important school in this field is the two-level games, presented by Robert Putnam. This theory suggests a close linkage between international negotiations and domestic support. Whatever the results on the international “table” (level 1) between two different teams (representing two states) are, they will be accepted and become operational only after they have been “ratified” in the respective countries by domestic procedures (level 2). An important concept here is that of winsets, indicating the sum of all possible results at table1.

Transnationalism

This concept perceives social actors as transnational (non-state) actors and players on the world politics arena. Their prospects to intrude other societies depend on those societies and their domestic structures.

Domestic actors have a „primacy.“ They act in a way that is goal-oriented, risk-averse and rational. Politics (including external) is the result of domestic bargaining. The basic social actors are relatively autonomous.

States and governments are not independent entities and actors, but representatives of different and particular interests. Those sub-state - interests aim at power and resources, via coalition building, elections and the media.

Critique: While generally quite plausible and strong in explanatory power, “liberal” approaches require research strategies which are complex. They may require case studies, and they are necessarily comparative.

You never succeed in identifying a domestic structure / coalition once and forever – they are permanently shifting.

Work group tasks:

1. What domestic structures are decisive for your region / conflict?
2. If you analyze the recent events in your region - is there anything which **cannot** be explained by domestic structures, constellations, interests?

COGNITIVE APPROACHES

This IR-theory is located at the first level of analysis. It may be said that it suggests a *homo psychologicus*.

What are cognitive approaches? They are „reflexive“, they „consider the role of policymakers' beliefs and images,“ and they challenge „much of Western thought and practice premised on the assumption of individual rationality“ (Rosati).

Given the media landscape as it is now, individual decision makers are often seen, or portrayed, as the central movers of politics. In that case, problematic personal histories may cause huge havoc in IR. The apparent presence of these cases and examples and their permanent availability makes alleged idiosyncrasies of decision-makers easy prey for the media.

Why are cognitive approaches needed or popular? The performance, and the behavior, of decision maker's, governments, bureaucrats etc. often seem difficult to explain in rational (goal-oriented) terms. The more complex an environment, or a task, or a challenge, the less „hard“ and reliable the facts and data.

Under strict time constraints, „psychological factors“ seem to gain in importance.

Basic assumptions of this approach are: Politicians act because of a specific social, ethnical, or cultural background. Politicians are, at least sometimes, characterized by (religious, ideological, other) obsessions. Or even, in some cases, politicians are „crazy“ and act irresponsibly.

Guiding questions of this approach: How do decision-makers view the world? What shapes their convictions and world views? But, first of all: What is the effect of decision-makers' perceptions and world views on IR/ WP?

Basic assumptions: The „subjective factor“ in decision-makers' reasoning is so dominant that it shapes their political actions. Other shaping factors are embedded into mindsets. Consequently, their actions cannot be explained by positivist approaches.

Cognitive approaches are interested in (seemingly) non-rational factors related to the mindsets of decision makers: particularities like world views, beliefs, specific decision-making habits and stress-coping potentials

Subgroups. Important directions for research in the framework of cognitive approaches are works related to images of the enemy / mirror images; the role of national stereotypes; the role of public opinion; and the role of perceptions and misperceptions. We will not describe these interpretations here in detail.

Operational codes: The concept of cognitive consistency is an overall coherent set of beliefs about the nature of political life. Among them, there are philosophical beliefs (on historical cycles, on the character of human nature etc.), and instrumental beliefs, related to strategies and tactic (risk-taking, etc.).

Cognitive mapping: This approach is related to specific beliefs, and their interconnectedness, in decision-makers' minds. It tries to represent a person's assertions about something, and it maintains there are stages – cognitive steps – of politicians' mental procedures.

Attribution (explanation) theory: This theory addresses attribution errors and biases. It tries to explain individual behavior with external / situational causes. It furthermore assumes strategies like accepting successes for oneself, blaming others for problems.

Social cognition and schemata theories: The basic assumption of this theory is that people do not like to change, once they have established beliefs / patterns. They screen their environment, and perceive data through their existing lenses. This is especially true in complex and unstable situations, and during informational overflow. The preferred strategy in such cases is Reductionism.

Some findings and results: It seems that there are strong indications that the organizational principle of cognitive consistency is important: „...individuals make

sense of the world by relying on key beliefs and strive to maintain consistency between their beliefs⁵.”

It also has been established that decision-makers try avoid (new) information not compatible with established assumptions / views and, therefore, requiring re-thinking. This makes common-sense assumptions such as „more information and proper consulting improve politics“ questionable.

Some research results imply that the effects of impacts, information, and propaganda are quite diverse, and strongly depend on contexts / environments.

Positive emotions toward „messengers“ and the personal surroundings are important. „Incoming information ... gets interpreted in accordance with an individual's existing central beliefs and predispositions.“ (Rosati)

Perceptions, patterns, idiosyncrasies, habits, exiting maps etc. are of utmost importance for the functioning of dm's, and for hb's in general.

Critique: New results from brain and neurological research imply that the reliability of our perceptions and memories is highly questionable. Similar problems with accuracy – itself a problematic concept – occur with our memories. Another crucial problem is operationalization. How can we reliably “open up the brains” of politicians and other actors to track their procedures?

Questions for work-grouping:

1. Look for obvious examples of dm's decisions that only may be explained by their *beliefs, operational codes* etc.
2. Name some typical belief systems of decision makers in your region.
How – if at all - could learning be initiated?
3. Imagine you are a newly appointed foreign minister.
What would you do to avoid cognitive traps?

⁵ Rosati, Jerel A.: A Cognitive Approach to the Study of Foreign Policy, in: Rosati, Jerel A. / Hagan, Joe D. / Sampson, Martin W. III (eds.): Foreign Policy Restructuring - How Governments Respond to Global Change. Columbia: University of South Carolina Press 1994, p. 52.

ISRAELI-PALESTINIAN CONFLICT

BY PETER DYLLICK-BRENNINGER AND CHRISTOF MAUERSBERGER

AN INTRODUCTION TO THE ISRAELI-PALASTINIAN CONFLICT⁶, FOLLOWING WALTZ' LEVELS OF ANALYSIS⁷.



Source: Central Intelligence Agency,
http://www.lib.utexas.edu/maps/middle_east_and_asia/israel_pol01.jpg

The conflict between Israelis (originally Zionists) and Palestinians over the land between the Jordan River and the Mediterranean Sea dates back to the early 20th century. Given this long history of the present dispute, a comprehensive introduction would provide enough material to fill book after book (some of those who have already been written can be found in the appendix). The task of our work group was to apply various theories of international relations to the conflict. Since these theories can easily be categorized by their level of analysis, the

following introduction will sketch out the most important actors on the three levels defined by Kenneth Waltz.

Here, only the actors relevant in the discussions of our workgroup will be introduced.

An important question that has to be discussed first is whether Palestine can be seen as a state or not. This is crucial, since most of the theories of international relations are actually theories of relations among states. Palestine is certainly not a state in a narrow sense. It has neither its own territory nor is it sovereign. However, it is still useful to work with the assumption that Palestine is a state for several reasons. The Palestinians are, at least partly, autonomous from Israel following the Oslo I (1993)

⁶ The information given here describes the situation as of summer 2004. Although the Israeli-Palestinian conflict is changing constantly, certain general patterns remain much the same. Especially the relations of actors and the structures on the various levels of analysis change slowly. This makes this introduction worth reading, even after a landmark event such as Arafat's death.

⁷ Waltz, Kenneth: Man, the State and War. A Theoretical Analysis, New York 1959

and II (1995) agreements. These treaties led to the establishment of the Palestinian Authority (PA) which ever since has represented and governed the Palestinian people living in Gaza and the West Bank. The PA is also negotiating with Israel eye to eye. This diplomatic practice seems to be the most important argument for taking Palestine as a state. Another one seems to be somewhat tautological: Most of the theories that we applied, and which do require states as actors, did produce sensible results. Considering Palestine as a state is therefore useful, but one should keep in mind that it is no more than an assumption: Palestine is not, at least for now, a proper state.

The *international level*, or third image, takes a look at the global context. This level can be divided into sub-groups: relevant third-party states, transnational actors and international organization; the relations between states; and the so called world system. The relevant states and other transnational actors are mainly those involved in the peace efforts (USA, Russia, Egypt, EU, UN and the Arab League) and those who are at least supposed to support Palestinian terrorism (Iran and Syria). Looking at the relevant relations between states, the most important one seems to be between Israel and the USA. Looking back on a long history of support, the USA stands firmly with Israel. A comparable partner, either by power or strength of support, cannot be found on the side of the Palestinians. The EU, and especially France as one of its most important members, seem to be strong supporters of the Palestinians. The Arab nations, however, are not as strongly on the Palestinian side as one might expect. Both Israelis and Palestinians consider themselves isolated from the rest of the world and as victims of the respective other side. From a systemic perspective, referring to the constellation of power, it is obvious that Israel is much closer to the center, while the Palestinians are closer to the periphery. Israel's military power and economic strength make it a natural regional center, although it is relatively isolated from its surroundings. The Palestinians in contrast do not have any military power, lack natural resources and are economically underdeveloped.

Most theories try to tackle international relations on the second or *state level*. The various approaches however use different perspectives on the state. Sub-national actors are at the focus of the liberal theories, the political system is the key for institutionalism and political culture and society play an important role in constructivist

approaches.

Who are the sub-national actors relevant to the conflict? The workgroup focused on the two major Israeli parties, Likud and Labor, the settlers' council and the peace movement for Israel. On the Palestinian side, we focused on the Fatah, the Palestinian Authority (PA), Hamas, and the Islamic Jihad as the most influential domestic actors. With regard to the political system, there is a clear difference between the two sides. Israel is a parliamentary democracy, which is marked by a wide variety of parties represented in the Knesset (parliament). This produces coalitions including several different parties, which in turn lead to unstable governments. The high frequency of national elections and coalition re-building is an excellent indicator for this instability. Although the PA, i.e. the ruling body for the Palestinians in Gaza and the West Bank, is supposed to be democratic, under the reign of Yassir Arafat it resembled a typical authoritarian regime. Given this structure, the PA government is highly stable. However, militant groups like Hamas or Islamic Jihad pose a credible threat to the rule of the established government from outside of the institutional setting.

From a social and cultural perspective, Israel and Palestine are just as different. The society of Israel is highly fragmented due to the varying origins of its citizens. Zionism and Judaism, as uniting ideologies, loose their importance as a result of the effects of the post-industrial era. The high number of Russian immigrants, who came to Israel mainly for financial reasons, increases this problem even further. On top of that, the Israeli government implemented harsh social reforms in the last years which increased social inequality. Economically, Israel is well positioned on the world market with leading companies in arms, information technology and biotech. However, the second intifada (starting in September 2000) did hurt this position. The Palestinian society is much less diversified than Israel's. It is, in spite of its difficult position, highly educated, yet the economical basis is still mainly agricultural, which is also a result of the occupation. The curfews, an everyday reality during the second intifada, hurt the Palestinian immensely. The hardships of the intensified occupation led to an increase in the importance of religion, i.e. Islam.

The first level looks at *individuals* who are relevant for international relations. On the Israeli side, Prime Minister Ariel Sharon is the most influential figure in the conflict, since he has the last word on all decisions. On the Palestinian side, Yassir Arafat has

an even stronger position (as of summer 2004). He is not only chairman of Fatah and the PLO, but most importantly President of the PA (Palestinian Authority). In contrast to Sharon, he does not need to fear defeat in elections. Another relevant figure for the Palestinians is Sheikh Yassin, the spiritual leader of Hamas. Given the importance of his party (or rather movement), he has a high moral standing and the power to influence public opinion and therefore politics by using this public opinion and by directing violence (i.e. suicide bombings). On the regional level, the most important individual is Hosni Mubarak, President of Egypt. Egypt was the first Arab country accepting Israel as a state, Mubarak himself pushed for a ceasefire on several occasions. Since he is respected by both sides, he plays a vital role in most negotiating efforts. From an international perspective, given the military, financial and political power of their nation, the most important person is certainly the President of the United States of America, Bill Clinton until January 2001 followed by George W. Bush.

Although this introduction is not only a snapshot of the situation of summer 2004, but also overtly brief, it should have made clear how different the two sides of the conflict are. For more information, we would like to suggest the following sources:

Web Resources:

<http://www.haaretz.com>

Online version of the English translation of Israel's number one quality newspaper. Very up to date.

<http://www.cnn.com/WORLD/meast/archive/>

Archive of CNN's Middle East coverage.

<http://www.nybooks.com/articles/14380>

Very interesting article on the negotiations at Camp David 2000, written by co-negotiators.

<http://www.merip.org/>

Website of the "Middle East Research and Information Project" - a think tank with a clear focus on the conflict. Excellent background essays.

<http://www.btselem.org/>

The Israeli Information Center for Human Rights in the Occupied Territories. Good source to get an impression of the situation in the territories.

<http://www.yale.edu/lawweb/avalon/mideast/mideast.htm>

A collection of all relevant legal documents relevant for the conflict (UN Resolutions, Treaties, etc.).

<http://www.memri.de>

Middle East Media Research Institute: Offers a newsletter in German/English with translated articles from mostly Arab media (newspapers).

<http://www.world-newspapers.com/palestine.html> and ***<http://www.palestine-net.com/news/>***

Two lists of Palestinian (online) newspapers.

INSTITUTIONALISM IN THE MIDDLE EAST CONFLICT

There are several regimes or institutions, formal and informal, relevant to the Israeli-Palestinian conflict⁸. We focused our discussion on the Roadmap, as an international regime dealing with our conflict. The "Performance-based Roadmap to a Permanent Two-State Solution to the Israeli-Palestinian Crisis" had been presented on September 17, 2002 by the middle-east quartet (including the USA, Russia, the UN and the EU). It put down three stages for achieving peace on the basis of a two-state solution. On the first stage, the Palestinians had to stop terrorism, reform the Palestinian Authority and hold elections, while the Israelis were to withdraw troops and freeze the building of settlements. On a second stage, a Palestinian state should have been founded, while an international peace conference should have been staged and international monitoring of compliance with the Roadmap should have been put into place. On the third and final stage, a final peace treaty should have been signed, solving the most explosive questions, such as the status of Jerusalem, the refugees and the settlements. As one might already know, stages two and three were never implemented and even the low-key stage one failed over time. However, looking back at this attempt seemed to us a good example for regime theory. The fundamental question our workgroup discussed was whether the Roadmap enabled or restricted the behavior of the parties involved.

⁸ Very important regimes for the two societies are certainly the Torah and the Koran, having a huge impact on the lives of the people. We decided against discussing this further, because those two regimes act primarily on their respective side and are not regimes mutually agreed upon by both parties.

First of all, we tried to grasp the features of the Roadmap making it an international regime. Every regime consists of a set of principles, norms, rules and decision-making procedures, and so does the Roadmap. Its principles are peace and security as the highest values, but also sovereignty of the people. The Roadmap furthermore sets non-violence, democracy and transparency as its norms. The rules of the regime are seen in the agreement to finally end terror, occupation and the building of settlements and to reform the Palestinian Authority. A decision-making procedure was designed and implemented by the middle-east quartet, i.e. USA, EU, Russia and the UN. Besides all these features, the Roadmap is neither an actor by itself nor has it an organizational body. So it does fulfil the criteria of a regime perfectly.

According to regime theory, the behaviour of the actors should converge around this setting. Consequently the question to be answered is about the Roadmap's impact on this conflict and how it worked. The most important aspect to any (successful) regime is the reduction of transaction costs for both parties. The Roadmap set common goals, defined the actors participating in the process and the decision-making procedures and freed the delegations to a great extent from discussing these matters over and over again. But still, critics argued that exactly those matters were not defined very precisely. For example, the first stage did not foresee a monitoring of compliance. However, the Roadmap supplied the two sides with information and therefore increased the willingness to cooperate, which is probably the most important aspect of this regime. It clearly defined *cooperation* and *defection* and hence stabilized expectations of both sides. Additionally, the pay-off-matrix was altered through this definition by increasing the costs of defection. This was the case, because both parties feared a loss of reputation, being disadvantageous in the upcoming negotiations and in talks on financial aid, being vital to both sides. The three stages laid out by the regime is a means to overcome the Prisoner's Dilemma by including the "shadow of the future". This means bringing the iterating character of the "game" to the attention of the actors and in doing so making cooperation the most efficient strategy of both parties. That is the reason why this regime was created and why it was useful.

Naturally, a discussion emerged whether the Roadmap was a "dead letter regime". Given the failure, it is an easy question to answer. Still, some steps on both sides have been taken. For example the office of a Palestinian prime minister was created,

thus giving Israel the opportunity to hold high-level talks with the Palestinians after the Sharon administration declared Arafat a “persona non grata”. Besides, a temporary unilateral truce by the Palestinians had been put into place and Israel did demolish some illegal settlement outposts in the context of the Roadmap. In the end, however, actual peace negotiations never took place, nor did the violence end. This might all be due to the noncommittal manner in which the Roadmap was written, since only the smallest common denominator seemed agreeable. Nevertheless, the effects discussed above are still relevant and true, even if only to a smaller extent. The best answer to all pessimism is still the counter-question: What would be the consequences if the regime would not have been put in place? Would it not have been even worse?

Analyzing only the Roadmap is not sufficient to conclude whether regimes play a decisive role in the Israeli-Palestinian conflict. As noted above there are various relevant regimes, formal and informal, both deepening and softening the conflict. It is important to point out that the introduction of formal institutions always brings about the replacement of informal ones. The failure of the road map might be read as a failure to replace certain informal institutions. So the reason why the Roadmap failed is not inherent in the concept of a regime, but rather evidence for a poor design. All the weaknesses are arguments to further develop measures increasing the will to cooperate rather than abolish regimes as such. A properly designed regime can enable and restrict the behaviour of conflicting parties and hence further the cause of peace.

FROM DOVE TO HAWK – CHANGES EXPLAINED

Domestic Structures/Liberalism

Theories of international relations that take state preferences as exogenously given often struggle to explain changes in foreign policy. By opening the „black box“ state and examining the domestic structures, liberal approaches are useful for understanding policy turns that are usually hard to grasp with third image rational choice models. To illustrate this advantage, our workgroup applied this theory to two questions: Why did the Palestinian leadership, i.e. Arafat, decide to support terrorism after the second Intifada started although he opposed it in the beginning? And why

did the Israelis decide to re-occupy territory during the second Intifada, although they had withdrawn from it years earlier? It is important to note that examining domestic structures is a very complex endeavor. The relevant actors and the configuration of power are constantly shifting, and hence using this approach makes a broad empirical basis necessary. Needless to say that our workgroup did not fulfill this requirement, although some preparatory efforts were taken. However, this naturally limited the depth of the discussion and of this article.

The domestic structures of "Palestine" can easily explain how the shift towards terrorism of the PA in early 2002 came about. Looking at the institutional level, the most obvious aspect is that the Palestinian Authority is *de facto* not democratic. Yassir Arafat reigns with autocratic powers, and there is no credible opposition against his government within these structures. The only opposition, strong enough to challenge Arafat's power, is located outside of any institutional setting: the Hamas. To understand the importance of Hamas, it is important to notice that it is much more than a terrorist organization. Its military wing is supported by a strong Islamic *social* movement, with strong roots in society. Especially in Gaza, Hamas is highly respected for providing social services for the poor and disadvantaged. As a consequence of this involvement, Hamas is in touch with the needs and interests of the people. The Palestinian leadership in Ramallah in contrast is quite detached from the masses. Arafat and most of the men surrounding him, spend the better part of their lives in exile, and did not return to the occupied territories until the peace process started in the early nineties. This absence, especially during the first Intifada, combined with a highly privileged living standard led to a detachment from the Palestinian public.

Now, at the onset of the second (al-aqsa-) Intifada all these factors became relevant. The masses started the uprising, instantaneously supported by Hamas and other militant groups. The Palestinian leadership, instead of taking a clear position, was waiting while the violence started to escalate and Hamas gained mass support for its suicide attacks on Israel. Under these circumstances, the PA was unable to judge how strong Hamas might become and if they might threaten its power, because there was no institutional setting, such as a functioning parliament, in which these differences might have been articulated. So, basically out of fear of losing power,

the Palestinian leadership decided to go along with the masses and support terror.

This policy can still be called rational if the level of political actors is our reference. Rational choice-theorists would call every decision of every actor rational, since it served their personal purpose and therefore was goal-oriented. On a collective level, this often leads to an irrational outcome, which might have been the case here.

Israel, a democratic state, made a similar policy shift, which, however, was brought about by elections. During the election campaign in late 2001 and early 2002, the second Intifada had already started. Ariel Sharon, candidate of the rightist Likud party, himself stimulated the uprising by visiting the “harm al-sharif”. This was perceived as a strong provocation for Palestinians, as they consider this holy site part of Palestine. Under the impression of rising violence and the failure of the peace process, the Israeli public shifted to the right and called for a tougher stand towards the Palestinians. Ariel Sharon won the elections and instantaneously translated this shift in public opinion into a much tougher policy towards the Palestinians, starting with “targeted killings” and culminating in “operation defensive shield” in spring 2002, meaning the forceful re-occupation of territory handed over to the PA under the Oslo-accords.

Policy changes in Israel can be much easier analyzed by looking at the institutional settings (i.e. majorities in parliament). Depending on which party gains more votes, policy shifts become possible; and: those shifts in foreign policy can best or even only be explained by looking inside the state. Explaining a policy shift within *one* government proves to be more difficult and would ask for a deeper analysis of Israel’s decision making process than just the look at parliamentary majorities. In our case however, this was not necessary.

The application of this theory to the conflict produced useful insights. Even though a lack of knowledge limited the discussion, many members of our workgroup mentioned gaining a deeper understanding of the changes in the conflict. An interesting point is that under this approach the push of the international community for a democratization of the PA makes perfect sense. Liberal approaches explain the empirical phenomenon that democracies never fight each other. Although such thoughts would not be more than speculation, it seems at least questionable that the

Palestinian people would have staged such an uprising when they felt fairly represented in their political system.

BARAK AND ARAFAT PLAYING (TWO-LEVEL-)GAMES

Camp-David-Negotiations Dec.2000⁹

In December 2000, during the final days of his presidency, U.S. president Bill Clinton pushed for an end of violence between Palestinians and Israel by getting their respective leaders to negotiate a final peace agreement. These long and intense talks became well known as the Camp David peace negotiations. As Putnam's two-level-game is a tool designed to model international negotiations, we tried to figure out if it would be useful in explaining the failure of these talks between Ehud Barak (prime minister of Israel) and Yassir Arafat (President of the Palestinian Authority). Primarily, we will take a look at the negotiations on the international level, in Putnam's terminology "level I negotiations". Since no ratification process took place (there was nothing to ratify), the domestic or second level will only be regarded indirectly through the win-sets, which are determined by domestic factors. As already mentioned in our introduction to the conflict, we regard Palestine as a state, given that Arafat was Barak's counterpart in the negotiations and had a separate level II. Representing the order of our discussion in the workgroup, a separate presentation of the Israeli and Palestinian win-sets will stand at the beginning. These presentations will be split in two sections, one describing the positions on the issues and another on the domestic factors influencing the size of the win-sets. Following this, a brief paragraph will discuss the outcome of the negotiations. The concluding paragraph will look into the usefulness of Putnam's model to explain the failure of "Camp David".

The major issues determining Arafat's win-set were the question of land and Israeli settlements, the status of Jerusalem and the question of the refugees. A crucial aspect for Arafat was the status of '67 was regarded as the starting point for the

⁹ Our main source was a very interesting article written by co-negotiators of the Camp David negotiations in December 2000: Agha, Hussein / Malley, Robert: Camp David: The Tragedy of Errors, in: New York Book Reviews, Vol. 48 No. 13 (August 9, 2001), <http://www.nybooks.com/articles/14380> (accessed on February 15, 2005).

negotiations and not the actual status quo. Therefore his win-set included a very high percentage of land to be returned, coming close to 100% of occupied territory (better: to be given *back* and not simply *given* as part of an Israeli offer). This can be illustrated by the fact that Palestinians see all of Israeli territory as originally Palestinian homeland; therefore the acceptance of an Israeli state is for them already a concession of land. Jerusalem as capital of a Palestinian state was a key point in Arafat's win-set. With respect to Arafat's claim on the refugee's right of return to Israel, there seemed to be greater leeway. Given these positions, it seems evident that Arafat's win-set was rather small.

Three additional aspects of the anticipated domestic "ratification-process" reduced the win-set further in size. First, past experiences in negotiations raised suspicion towards any Israeli proposal. Given the perceived history of Israeli defections from past promises, Arafat's greatest concern was not to give the impression to domestic veto-players of being deceived by Israel once again. Second, Arafat's autocratic position left the militant islamists as veto-players being the only obstacles in the ratification process. The extremists are articulating their opposition not in terms of votes, but rather through their influence on public opinion taking people to the streets and perpetrating violent attacks on both sides of the green line. Hence an unfavorable outcome could have sparked a Palestinian civil war. Finally, the costs of no-agreement for the Palestinian President were negligible, as his political position in the given situation was not at all threatened in case of a failure; a poor agreement would have been much more risky. As already mentioned, the basis for the negotiations for Arafat was the status of 1967 rather than the status quo. This further lowered his perceived costs of no-agreement, as he also expected the position of Palestine to improve in the long run. International pressure on Israel to give up occupation would rise, as would economic pressure in Israel and the demographic conditions of Arabic population. It could also be assumed that Arafat had the history of Lebanon in mind, where Hezbollah achieved Israeli withdrawal over time by use of force only, not making any concessions. All these facts pushed the costs of no-agreement for Arafat to zero and he therefore neglected the need to enlarge his win-set.

Barak's primary objective was to reach a final agreement, settling all issues between the two sides, rather than a step-by-step approach. His attitude was marked by distrust towards the Palestinian leadership. The core issue for the Israeli side was

security. Settlements in the West Bank and Gaza were seen as Israeli territory, but following the “land for peace”-doctrine, Barak was willing to give up most of this land in order to accomplish a peace-settlement. The notion of Jerusalem as part of a Palestinian state or even as its capital was rather unthinkable for Barak. A comparably strong position was taken on the issue of the refugee’s right of return to Israel, as a probable influx of Palestinian refugees into the Israeli heartland was feared to end the Jewish character of Israel. What clearly enlarged the win-set of Barak was the great demand of the Israeli public for a final agreement providing security and stability, i.e. peace.

Clearly distinctive in determining Barak’s win-set was the institutionalized and thus predictable ratification process on the Israeli side through the Knesset. Given party majorities in the parliament and favorable national polls made it easier for the prime minister to calculate his win-sets. The missing predictability of Palestinian ratification process made it more difficult for Barak to assess his counter-part’s win-set. The Prime Minister also couldn’t be sure whether an agreed contract would be followed also by radical Palestinian opposition groups, since they are not integrated into the structures of political power in the Palestine Authority but articulating their protest “on the streets”.

It seems impossible, at least with the information at hand, to exactly determine the win-sets of both sides. Even though an agreement on the issues seemed to be possible, the win-sets finally didn’t overlap. This was primarily due to the Palestinian domestic constellations analyzed above, which put Arafat in a position where he didn’t feel able to make any concessions. Additionally, the lack of knowledge about each other’s win-sets seemed to be fatal and consequently led to serious mistrust. Arafat was most concerned not to reach an agreement, but to be seen as a tough negotiator defending Palestinian interests and not to be fooled by the Israelis. The perceived low costs of no-agreement by Arafat also contributed to the failure of Camp David. The existence of silent allies on the respective other side should have made the peace talks easier, but there were also allies opposing peace on both sides.

Putnam’s two-level-games could be well applied to our case and certainly proved useful to understand the negotiations. As it is often the case with models, the problems in application lie in the details, here especially in determining the exact composition of the win-sets. The Two-Level-Games are therefore to be seen as an

instrument to grasp the dynamics of such international negotiations, their links to the domestic area and the motivation of the chief negotiators (which we didn't really inspect here). They are a very good tool to connect international and domestic level and to study their links and dynamics, rather than to examine both levels separately.

COGNITIVE APPROACHES IN THE MIDDLE EAST CONFLICT

Decision Maker: Brain

Cognitive Approach focuses on individual decision makers as processors and managers of information. Consequently, states are neither unitary actors nor actors at all. Decisions are taken at the individual level and hence the individual's belief system must be examined to understand how. We will therefore take a closer look at the two most important decision makers of the conflict, Yassir Arafat and Ariel Sharon, and their belief systems. Afterwards, we will try to illustrate the impact of these belief systems on particular decisions. Before evaluating the benefits of a cognitive approach for the analysis of the Middle East conflict, we will examine the prospects for initiating a learning process.

Arafat has a military or rather: guerrilla background. He has lived most of his life underground or in forced exile before returning to Palestine and becoming the political leader of his people. He repeatedly experienced the use of violence as helpful in achieving political goals. His self-image of a fighter is expressed by the battle-dress he was constantly wearing, even when meeting other statesmen. Arafat is said to have some kind of paranoia of losing his power by betrayers inside his own ranks, so he is very careful not to enrage any potential opposition forces. This narrows his scope in negotiations and makes him look almost exclusively at domestic constellations by ignoring Israel's restrictions. Besides these personal aspects, cultural norms and values also play a vital role as impersonal social powers. In Arab societies, defeat in general is seen as maybe the biggest disgrace or humiliation. Strength and power are therefore central values, so that Arafat has to present himself as a strong defender of Palestine rather than as somebody who is able to agree to compromises. The historical background is especially important in explaining Arafat's, and Palestinians in general, very strong self-perception as victims of Israel,

but also as victims of Arab neighbours and as victims of the super powers, who all neglected their support for the Palestinian struggle for independence. Given the many incidents of suppression by Israel and Israeli defection from treaties, Israel is not seen as a partner in negotiations, but rather as an enemy against whom one has to defend one's own interests.

Although coming from a very different cultural background, Sharon's belief system shows some similar features. His career started at the military, where he spent almost his entire life. The ex-General fought in all of Israel's wars since 1948 and it therefore seems obvious that he has no fundamental objections towards the use of military force to achieve political goals. He rather fears that weakness might lead to the destruction of Israel. For him, the Arab slogan, "Throw the Jews into the sea" is still representing a credible menace. In the eyes of Sharon, this existentialistic threat can only be met by military force, since Israel and the Jews never got any support from outside forces (Holocaust-experience). Sharon also enjoys the image of being a political underdog in the political scene in Israel. He always has to fight opposition from the leftist political establishment. Standing against opposition seems to make him even stronger and more determined to have it his way. This makes him, just like Arafat, resistant to compromises. Those features of Sharon's belief system already indicate some difficulties that might arise when negotiating the Palestinian future.

Belief systems are filters and thus do not *determine* any actions. Decisions are still taken i.e. with regard to military/political restrictions and to the domestic public, but may be *altered* by personal aspects of the decision maker. Therefore it seems hard or even impossible to identify any specific decisions which can only be explained by cognitive aspects, but still: Other outcomes are possible, if the same circumstances are perceived through a different belief system. The most obvious influence in our case would be the military experience of both Arafat and Sharon, who therefore seem to "filter out" political options and swing to military solutions. In the face of the rising second Intifada, both decision makers chose violent paths – Arafat by supporting terror and Sharon by forcefully reoccupying the territories already administered by the PA. Arafat left the path paved by the Oslo peace process and returned to violence by supporting the second Intifada in early 2001, while Sharon answered with the "Operation Defensive Shield" rather than diplomatic solutions. Both leaders distrust

political solutions, which can be explained by their biographical experience and not necessarily by assuming “rational”, goal oriented behaviour. Another example for the influence of cognitive aspects on the conflict is the refusal of Sharon to meet or talk to Arafat. Sharon has seen him as a terrorist and his personal enemy ever since their military encounter in Beirut, Lebanon in 1982. The cognitive approach stresses the importance of inter-subjectivity for the constitution of reality. Here, this would mean, that since they both consider each other as leaders prone to use violence, in reality violence is being used and regarded as the only option.

An important question for cognitive approaches is the possibility of learning. Can learning be initiated, and if so, how? With regard to Sharon and Arafat, this must be seen as rather difficult. Their advanced age, along with a lot of political experience, is certainly the most telling indicator hinting towards a lack of learning capacity and the difficulty to change their perceptions and beliefs. Both of them appear to be examples of cognitive consistency, since they seem to ignore any information that does not fit into their belief system or make this information fit in by misinterpreting it. A good example is the failure of the Middle East Quartet in changing their policies. This can be interpreted as a failure in changing Arafat's and Sharon's perceptions of reality which is in fact filtered by their belief systems. Occasions that, according to theorists, induce learning, such as important personal, or societal events, did not occur while Sharon and Arafat were in power.

The cognitive approach seems to give some helpful hints towards understanding the situation as of summer 2004, especially the lack of movement. But it is impossible to conclude that the choices made can only be explained by the decision makers' belief systems when it comes to specific decisions. There are always political/military “realities” that restrain behaviour and there is always a public opinion that influences the decision making process. As mentioned above, belief systems function as a filter rather than a source for decisions taken. From this perspective, the emphasis on the individual leader seems to be productive. It opens up the possibility of changing the politics of a state (be it by changing the leader) and makes it possible to include personal features of decision makers. It would seem absurd to deny that personal experience and cognitive perception of human beings (which politicians are, too) would not have any effect on their decisions taken. A more Gandhi-like political

leader in the Middle East would definitely have contributed to a different outcome in the region, even if the military, political and historical circumstances would have been the same.

IRAQ CONFLICT

BY KATHARINA HOFFMANN

CONFLICT SETTING: IRAQ – AN INTRODUCTION



Source: Central Intelligence Agency,
http://www.lib.utexas.edu/maps/middle_east_and_asia/iraq_pol_2004.jpg

When elaborating core elements of a long-lasting conflict, one is confronted not only with the most influential agents, but also with several structural aspects, determining the conflict situation. Depending on the respective theoretical approach, geographical features, deposits of natural resources, historical developments (or rather their instrumentalization in the conflict), the ethnic, religious and social structure of society, may serve as explanatory factors.

In order not to overstress our analysis, on the most recent crisis of 2002/2003, which came to a head with the claim of US President George W. Bush to enforce the destruction of weapons of mass destruction, upon which the UN Security Council Resolution 687 (cease fire resolution 1991) had agreed, and Iraq's rejection to cooperate with the inspectors of UNMOVIC¹⁰ and IAEA¹¹, which according to the UN Security Council Resolution 687 and 1441 (2002) had the mandate to monitor the disarmament. This confrontation led to the third Gulf War. The US-led operation "Iraq Freedom" started with air raids on Baghdad on 20 March 2003 and ended on 9 April after coalition tanks rolled into Baghdad.

Since the aim of our working group primarily lies in testing the explanatory power of different theories of International Relations and not in analyzing causes and solution strategies of the conflict, this introduction will present a relatively reductionist outline of the main agents involved in the conflict. Good theories, too, are in principle parsimonious and simplifying.

¹⁰ UNMOVIC: UN Monitoring, Verification and Inspection Commission

¹¹ IAEA: International Atom Energy Agency

We will hence refer to Kenneth Waltz' three "levels of analysis"¹² providing fundamental tools of analysis in International Relations. The first level stresses the role of individuals as agents, whilst the second image is focusing on the societal and national context. The third image is related to the global structure of the international system, formed by coactions of the units. Therefore the relation between states and relevant transnational and international actors have to be taken into account.

The Iraq Crisis is not an internal conflict, so it appears only natural to first concentrate on the third level. One of the most important international actors are the United Nations and the IAEA. In 1991 the Second Gulf War following Iraq's invasion in Kuwait ended with the UN Ceasefire Resolution 687, imposing the destruction of biological, nuclear and chemical weapons under control of a UN Special Commission (UNSCOM). Due to its position to provide an international legitimacy for military actions against Iraq, the United Nations had functioned more or less as a mediator at least until March 2003. The UNMOVIC and IAEA had been in charge of verification for Iraq's deposits of BNC weapons and provided substantial information influencing the decision making process of the UN members. Regarding the relations between states most important seem to be the relation between Iraq and the USA as well as Iraq's attitude towards the UN. Iraq's cooperation with UNSCOM had been characterized by diverse attempts to impede the inspector's work. The inspections had been almost totally blocked from 1997 to November 2002, being resumed only in reaction to massive international pressure. The tension between the USA and Iraq can be assessed as critical, since the USA still perceived the Iraq as a threat to international security. Applying the classification of the world system approach the Iraq belongs to the periphery being isolated from international economic and political cooperation due to the sanctions stipulated in 1990 (UNSR 661), while the USA constitutes one core center of the world system.

A look on the second image including national and sub-national actors may be illuminative concerning the emergence of 2002/2003 crisis.

The United States Government, supported by 30 states which formed the "coalition of the willing", appears to be (aggressor in) the initiator of the third Gulf War, arguing that Iraq's possible use of weapons of mass destruction necessitates "pre-emptive action". In September 2002 U.S. President George W. Bush raised the issue of Iraq's disarmament in front of the UN General Assembly, declaring that should the UN

Security Council not enforce the process, the United States would consider unilateral strategies of disarmament. There are a variety of interests of the U.S. Government in the conflict including the “war on terrorism” proclaimed in consequence of 9-11, internal legitimacy, as well as geo-strategic and economic aspects concerning e.g. the deposits of oil in the region. Turning to the U.S.-American sub-national level, public support of the Government’s Iraq policy had been generally high. Focusing mainly on the question of unilateral action versus multilateral action, two camps within the Bush administration could be distinguished: On the one hand, the neo-conservatives (represented by Vice President Richard Cheney as well as Defense Secretary Donald Rumsfeld) perceived Iraq as an immediate threat for vital national interests and believed that an involvement of the United Nations would only delay necessary measures. On the other hand, there was the assumption (by Secretary of State Collin Powell et. al.), that unilateral action would damage U.S. long term interests.

The debate on the strategy to deal with Iraq divided not only the UN Security Council but Europe as well. Whilst the leaders of Great Britain, Denmark, Italy, Portugal and Spain together with most of the then EU candidates supported of the U.S. policy towards Iraq, Germany and France et.al opposed unilateral military action against it.

The United Kingdom, due to the important role it played in the history of modern Iraq, can be described as the traditional and most important partner in the US-led “coalition of the willing” concerning US-Iraq-policy. In spite of backing the US position in the course of Gulf War III in official statements very early, the decision to participate in military action even without an UN Resolution had been made only after the USA had already given Hussayn the ultimatum.

The neighboring countries, even though most affected by direct consequences of a war, could exert only limited influence on the question of military action against Iraq. Apart from Jordan, which sought not to jeopardize its close relations to Washington, the regional actors rejected the U.S. policy towards Iraq with differing intensity. Different motivations, as et al. the fear of Iraq’s disintegration leading to regional instability or an interest in keeping U.S. influence in the region at bay are decisive. In view on Iraq the main actor on this level appeared to be the Baathist regime. The Baathist regime in Iraq could be described as a dictatorship with strong features of a patronage system, relying on tribal structures, domestic repression and economic inducement. Apart from the Baathist Party and the bureaucracy, the armed forces

and security services, being split in several competing organizations, are said to have been a main source of regime stability in Iraq. Broadening the range of options for Hussayn by supporting the regime they play a vital role in the conflict. The organized opposition in exile, though, plays an important role. The almost 40 parties and religious movements are taking action from their London base and gain vital political support from Washington. In scope of the "Future of Iraq Project", which has been initiated by the U.S. Government, the Iraq National Congress and the Supreme Council of Islamic Revolution in Iraq, elaborated, a strategy for transition to democracy. Furthermore the Iraq interim Government had been formed out of members of these opposition parties.

Finally the individual level has to be scrutinized. The two main characters immediately coming to mind are Saddam Hussayn and George W. Bush. The excessive power of Saddam Hussayn arose from the system of power he had established, since it was based on personal ties and loyalty to his person. Regarding George W. Bush there might be mentioned, that his politics were determined by his conservative attitude and strong religious back ground.

As mentioned before, this introduction is supposed to give only a short overview of the conflict setting, in order to provide a first fundament for the debates in the working groups. Below you will find some suggestions for literature and interesting web sites for a more in-depth analysis of different aspects concerning the situation in Iraq.

Further Readings:

Fürtig, Henner: Kleine Geschichte des Irak. Beck, München 2003.

Krause, Joachim: Die Krise um den Irak und die internationale Ordnung, in: Kieler Analysen zur Sicherheitspolitik No.4, Jan. 2003, p. 1-25.

Lawrence, Christopher: Iraq, 9/11, and the War. Understanding Mass Belief in the Perceived Threat of Saddam Hussein and Support for War. Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago Illinois 2004, p.1-15.

Münkler, Herfried: Der neue Golfkrieg. Hamburg: Rowohlt 2003.

Press-Barnathan, Galia: The War against Iraq and International Order: From Bull to Bush, in: International Studies Review, Vol. 6. 2004, p. 195-212.

Http:// www.state.gov US Department of State: Official statements

Http://www.rferl.org Radio Free Europe – Radio Liberty

Http:// www.icg.org International Crisis Group

DO IDEAS MATTER?

A Constructivist Approach To The Iraq Crisis

What does the 2002/2003 Iraq crisis look like from a constructivist point of view? After having a short look on texts and essays introducing constructivism in IR, the question upcoming first is: Does “the” one constructivist point of view exist and does it provide a clear frame of laws and hypotheses, which we can apply to our case? Constructivism in International Relations rather implies a wide range of approaches, which propose constructivism as a meta-theory, a social theory or theoretical and empirical perspective, rather than a clearly shaped homogenous concept. Nevertheless, there are some common assumptions upon which the different branches are based.

First, the ontological perspective assumes that reality is socially constructed and cannot per se be experienced. It can be assessed, or “constructed” but only against the background of discourses in which the agents and observer are integrated. As a result the epistemological assumption considers knowledge as a construct as well. This leads to the question how these social constructs are produced and which discourses are decisive for certain settings.

Since this provides more of a breeding ground for a (even philosophical) debate which easily gets out of hand, and therefore is not suitable for a short discussion on crisis in Iraq, we decided to refer to Alexander Wendt’s Social Theory of International Politics.

Wendt concentrates on the problem how preferences and interests of actors can change. According to his theory one core element of analysis is the identity of the agents. Identity as well as interests, deriving from identity, are based on shared ideas and are socially constructed in course of interaction between different actors. Though he does not completely evade opening the black-box “state”, Alexander Wendt mainly looks at the third level, because he aims at explaining structural changes in the international system. He as well proceeds from the assumption, that the international system is principally anarchic, but points out that it is structured by identities, depending on social and individual cognition. The social set which can be found in the international system ranges from adversary, rival to friend.

Against this background, the following questions structured our discussion:

Which patterns of identity and interests had been decisive in the Iraq crisis? Which ideas and perceptions had been significant for the change of international cooperation in course of Iraq crisis and before? Does a re-definition of identity can be observed?

In order to get an idea of the structure and the perception of roles determining the international system in the period of the conflict, we tried to find out patterns of identity of Iraq and USA, while first focusing on the mutual perception. In the second step we concentrated on the countries' self-perception. The mutual image of both USA and Iraq had been more or less the same, only the argumentation differed. They considered each other as enemies. That implies a high willingness to use violence against each other. As officially argued by the USA, Iraq presented a threat toward international security and therefore a threat to the U.S. – American security. From the Iraq point of view, the United States did not only undermine the Baath' regime through various action, but threatened it directly by claiming a regime change in Iraq. Due to the fact that a digression from the structure of the international system would be too far reaching, it should only be mentioned, that the U.S. assessment of Iraq's role had been supported widely. In contrast, the role attributed to the U.S.A. had been mainly the one of a "friend". Since this set of perceptions can, at least after the second gulf war, be considered as a relatively stable, we argued that the mutual attitude has only limited importance for the emergence of the current crisis.

Discussing elements of Iraq's self-perception, we figured that the claim to become or to be a regional power had been decisive for Iraq's foreign politics, but at present at least no action in this direction can be observed. In view of Saddam Hussein's non-cooperative position, one can assume that the stability of his regime was assured (beyond doubt) in his own perception. To sum up it can be said that the self-perception of Iraq had been relatively stable as well.

Investigating the U.S. foreign policy, some changes being decisive for the crisis, can be determined. Most obviously is the decreasing importance of international cooperation in U.S. foreign politics, since the military action against Iraq without U.N. approval.

Changes in means, changes in interests, changes in identity?

A main feature of U.S.-American self-perception is the image of a world power by means as well as by mission. Moreover, self-description as democratic society based

on Christian values, constitutes a significant characteristic of U.S.-American identity, which in recent times serves as criteria for negative identification against others.

The concept of democratic peace as well as the assumption that democracy and stability are exportable runs through their foreign policy strategies. An important discourse broaches the issue on strength and, coming along with it, security. The image of the invulnerable society has been deeply challenged by the Al-Quaida attack on the World Trade Center in 2001. 9-11 appears to be a turning point, because it triggered a change in the discourses concerning U.S.-American identity. Efforts to deepen the surveillance of potentially dangerous tendencies at the level of society and to establish an alert system for terror attacks give evidence for a profound perception of a security deficit.

As mentioned above, the division between societies based on Islamic values and Christian societies became a frequently used rhetoric. In scope of these recent processes the re-definition of Iraq as an imminent threat and the new consideration of means and options can be explained as an attempt to regain the feeling of security in the U.S.-American society. Self-confidence and security can be seen both as socially constructed phenomena. Because of their reflexive character the revaluation of the perception can result in consequent and successful action against a clearly defined opponent. Since Al-Quaida has not the qualities of an equal partner, being neither state nor international organization based on international law, which in consequence puts constraints to direct sanctions, the USA tried to shift the conflict to the state's level, while taking action against states potentially protecting terrorist groups. The fact, that the situation in Iraq had been a repeated topic since the war in 1991, can be seen as a supportive element for the decision to concentrate on Iraq.

As already mentioned, a revaluation of cooperation can be observed in the U.S. Foreign Policy during the crisis. This rises the question whether the change can be explained as well by going back to the discourse of security? According to Wendt, cooperation becomes likely if interests and actions of states are based on shared ideas. In course of the military action against Iraq, the United States did not principally deny cooperation, but the strategies mapped out by the U.N. did not match their U.S.-American objectives. Arguing that the first priority of the U.S. strategy has been to regain the feeling of security, it stands to reason that the interests differed from those of the most international actors. A continuation of U.N. inspections and sanctions leads, if at all, only to long term success. It could have been regarded

rather as weakness, than as an expression of the power and capacity to act, as it had been considered to be necessary for reestablishing the tarnished image of the countries security.

The decision to engage in military action with disregard to the U.N. position had been supported by other ongoing processes and discourses, like the softening of the “sovereignty” concept as institution in the international system. An increasingly intensive debate on humanitarian intervention in the early 1990ies, which has been closely linked to the war in Bosnia and Herzegovina, led to a widespread acceptance, that in case of doubt humanitarian intervention overrides the violation of sovereignty. This example shows that transaction cost can be reduced by such changing perceptions.

Albeit our analysis had been without doubt simplifying, not only by taking exclusively U.S.A. and Iraq in consideration and arguing mainly at the third level, the constructivist approach proved to be helpful for explaining and understanding key processes of the Iraq crisis in 2002/03.

DECISION OVER A “SECOND RESOLUTION” - A PRISONER’S DILEMMA?

Rational Choice Approach

After the introduction to Rational Choice as a meta-theory our working group had been asked to apply the prisoner’s dilemma – one of the most popular games of game theory – to a situation taken out of the Iraq crisis and to discuss afterwards the explanatory power of this approach.

In spite of the parsimony of this theoretical approach, the question, which situation would fit in the game’s setting had not been easy to answer. Since the model promises explanatory power for negotiations, we decided to apply the PD-model to the question, whether to pass a “second resolution” in the security council, which would authorize the United States and the so-called “Coalition of the willing” to implement Iraq’s disarmament via military campaign. Following a nearly 3 month debate, Great Britain, Spain and the United States in February 2003 came forward with a draft resolution, which declared Iraq’s breach of UN Security Council Resolution (UNSCR) 1441 and referred to Chapter VII of the UN Charter, considering

action in the case of threat to the international peace and security. France, Russia and China, three of the five veto-holding members of the Security Council, voted against military action, preferring an extension of the restarted UN weapons inspections. Hence the resolution failed, which did not prevent a military campaign against Iraq.

Just as the setting of the prisoner's dilemma is based on a short story, we first tried to develop the pursuant "story" for our setting.

Therefore we had to outline the interests of the "prisoners," who in this case are the United States, Great Britain and Spain (***prisoner 1***) on the one hand and Germany, France and China (***prisoner 2***) on the other hand. The overall common interest seems to be international security and peace and as a result the disarmament of Iraq. The best practice option for ***prisoner 1*** would be military action against Iraq legitimized by a UNSCR. For ***prisoner 2*** the best option would be to avoid military action in order to maintain international stability. Cooperation would mean the adoption of a common UN resolution resulting in a common strategy towards Iraq. Defection therefore would be rejection of a common resolution resulting in unilateral action. Four situations are to be considered:

Situation 1: The "coalition of the willing" agrees to continue the weapons inspections. In the case that evidence is discovered for Iraq's arsenal of weapons of mass destruction (WMD) they will argue that "pre-emptive" military action is legitimate given the right to self-defense defined in Article 51 of the UN Charter. The second resolution draft passes, providing the mandate for military action, if WMD are really found. According to the "prisoner's dilemma", this scenario should provide high costs for ***prisoner 1*** and low costs for ***prisoner 2***. In our setting the cost for Germany, France and China and for the UN are at least lower, in the case that the draft resolution passes – given the evidence of a material breach of UNSCR 1441 on the part of Iraq. Although military action is generally undesirably, collective military intervention within the framework of the UN is perceived as not to jeopardize the stability of the international order. Furthermore the operation draws legitimacy from the fact that Saddam Hussein denies compliance with international demands and presents a threat to international security.

From the US's point of view, the costs of collective action are high, since this would come along with further attempts to resolve the crisis with political means and the continuation of the UN inspector's work, which already proved to be unable to realize

their mission. This loss of time causes a highly risky situation concerning international peace and would not match the US interests.

Situation 2: Arguing that Iraq poses no imminent threat, the UN Security Council rejects a second resolution. Hereupon the “coalition of the willing” decides on military action against Iraq without UN approval. The theoretical game setting entails very high costs for “*prisoner 2*” but complete success for “*prisoner 1*”. Within the scope of our story, this would mean more or less success for the United States and its allies, because they assume to disarm Iraq with their preferred means. The costs for Germany, France and China are very high, because they are not able to prevent unilateral military action and aggression against another state and thus have failed to solve an international crisis within the framework provided by the UN.

Situation 3: According to the PD model the third constellation should entail high cost for each of the prisoners. In our setting the cost for both prisoners would be relatively high if the resolution passes and both sides would realize a military operation against Iraq on the basis of only vague evidence concerning Iraq’s WMD. Even if the operation leads to complete disarmament of the regime, the lack of fundamental evidence which is given concerning the topic of WMD in the country, would question whether Security Council Resolutions are able to ensure legality in the international law of military action any more. Referring to this argument, some of the UN members would not support military action, even if it would be authorized by a UNSCR. This would suggest that the United States would face complications in diplomatic relations with opponent states like Germany or France. Nevertheless, compared with unilateral action an UN approval would lower US-American costs. Last but not least there are the real costs of war and the country’s reconstruction in the aftermath of war.

Situation 4:

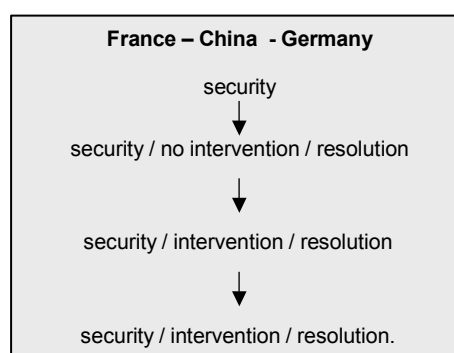
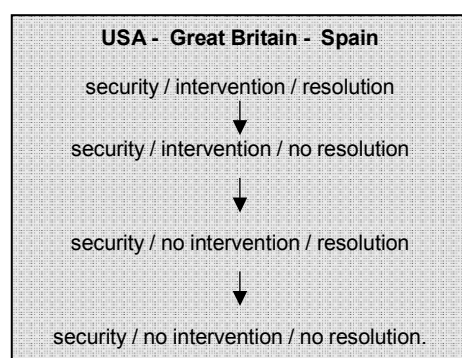
The less cost-intensive setting for both would be the solution of the crisis without military action. Taking into account the interests of the USA and Iraq, it appears difficult to say what such a solution should look like, but one proposal would be proven disarmament of the regime in Iraq reached by peaceful means, such as the work of UN inspectors and political negotiations. Since this setting does not depend on the decision of the two prisoners only, but on the degree of cooperation accepted by Iraq, it does not fit the PD model very well.

MATRIX OF THE CLASSIC PD MODEL AND THE MATRIX OF OUR SETTING

Prisoner 2	Prisoner 1		
		Not to confess (cooperate)	To confess (not to cooperate)
	Not to confess (cooperate)	5	10
	To confess (not to cooperate)	5	-10
		-10	-5
		10	-5

China, France, Germany (Prisoner 2)	USA, Great Britain, Spain (Prisoner 1)		
		Resolution and common strategy towards Iraq (cooperation)	No Resolution and no common strategy towards Iraq (defection)
	Resolution and common strategy towards Iraq (cooperation)	2	6
	No Resolution and no common strategy towards Iraq (defection)	10	(unilateral military action) -10
		-2	-5
		10	(unilateral military action) -5

SET OF PREFERENCES



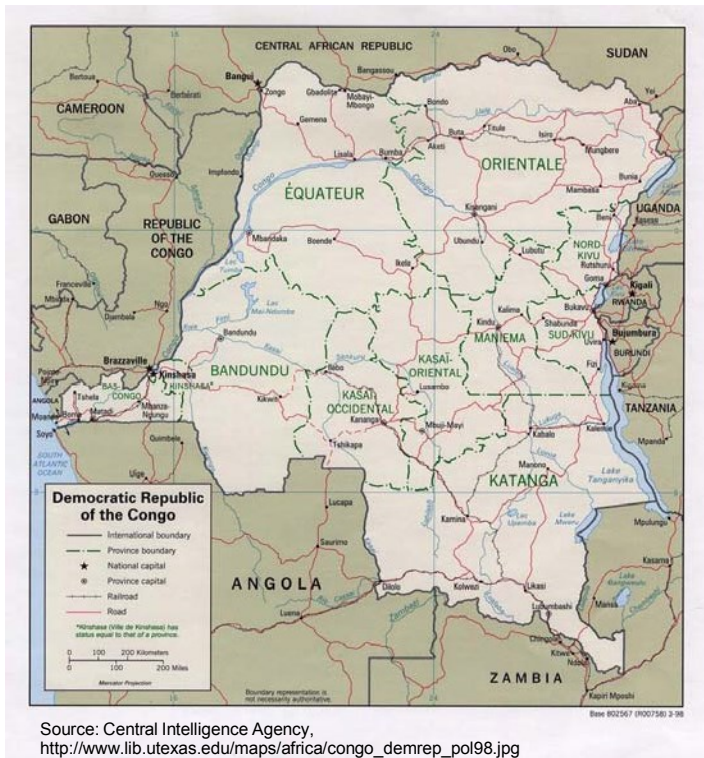
After all, does the setting developed for the second resolution correspond to the PD model and thus can the outcome be explained by it? Having a closer look at the four constellations, first of all, the argumentation obviously is not stringent. When comparing the original PD matrix and the matrix developed within the working group, which gives an overview over the possible settings and the respective costs, it is obvious that the cost for the prisoners in at least two settings differ from the model. The differences are caused by several preconditions which differ in the chosen example. First of all the situation focusing the UN Security Council Resolution is, of course, more complex, than the situation assumed in the model. Rational Choice theory requires a stable and straight set of preferences which lead to action. Our example appears to be too complex to reduce the actor's preferences to a hierarchical set of four or five preferences. While analyzing the interests of the UN we have to consider the special interests of the veto-holding members. Obviously, they cannot easily be reduced to a consistent interest; even the interests and preferences

of the USA cannot clearly be reduced to one strategy. They depend on several variables such as support within society, economic development, and the international position of the respective (sub-) actors, to mention only the most evident factors. The set of preferences of both prisoners in our model are varying (see table 2). One of the highest preferences of the “coalition of the willing” seems to be intervention in order to establish security. France, Germany, and China in contrast highly prefer the avoidance of intervention. Furthermore, the external setting cannot be considered as clearly determined in the case of the “second resolution” as opposed to the model, because there is no district attorney or officer determining the punishment. According to the PD model the prisoners both would not cooperate and therefore both have to face relatively high costs. The final decision of the UN Security Council regarding military action against the Baghdad regime had been not to pass the resolution draft. The “coalition of the willing” then started the operation “Iraq Freedom” without UN-authorization. So the agents did not choose the expected outcome, but the United Nations (France, Germany and China) had to assume a very high cost, while the cost for the United States were said to be relatively low in terms of pursuing their interests (see setting 2). Evidently the explanatory power of the prisoner’s dilemma had failed. The question remains, what are the reasons for this failure? Comparing the basic assumption of the prisoner’s dilemma with their application to our example, the agents’ perception of the external setting can be identified as remarkably different. While the strategies of both prisoners are based on the same information about the external environment and options, which had been explicitly pointed out by the district attorney, the agents’ perception of the external setting in the Iraq crisis differed widely. France, Germany and China, relying upon the results of the UN inspections, did not see the necessity for urgent action; whereas the United States perceived the Iraq Regime as an imminent threat for the national interests as well as international security. Given the deviating perceptions, the set of preferences differed to a degree that made cooperation unlikely. During the final discussion, the explanatory power of the prisoner’s dilemma for our example had been doubted. The working group argued that many of the core assumptions the setting relied upon had to be further scrutinized. One of the questions which arose is whether the disarmament of the Iraq had been the U.S. Government’s prior aim, or whether it primarily functioned as an intermediate stage for another strategy. In this case the set of preferences would be respectively altered.

THE CONFLICT IN THE DEMOCRATIC REPUBLIC OF CONGO

BY CHRISTINE SCHUSTER:

“THE CONFLICT OF THE GREAT LAKES”¹³



This text is supposed to give an introduction into the Conflict in and around the Democratic Republic of Congo. After first of all indicating the extent of the conflict with some figures that make clear why it's called war, a very short historical introduction is given, comprising the colonial time up to today. In this historical overview the most important actors and the essential sources of conflict are named, as are the most effective (if at all) peace steps. The introduction ends

with an exclusive image of the actual status and the future problems.

The Conflict in the DRC (formerly known as Zaire) has been called Africa's First World War There have been a number of complex reasons for the designation of the conflict, including competition for basic resources such as water, access and control over rich minerals and other natural resources which can be found in this central-African country that is the size of half of Europe. This led to various political disputes, especially in the two richest provinces, North- and Southkivu.

Since the outbreak of fighting in August 1998, at least 3.3 million people, mostly women, children and the elderly, are estimated to have died because of the conflict,

¹³ „The Conflict of the Great Lakes“ is a literal denomination for the war in and around the today Democratic Republic of Congo, former Belgish-Congo, Republic of Congo and Zaire. It refers to the rich areas around Lakes Tanganyka, Albert and the flow-ins of Lake Victoria, representing the ending of the great river Congo. This area has always been and still is a source of conflict about exploitation rights and political control.

most from disease and starvation. In addition to that, more than 2.25 million people have been driven from their homes, many of them beyond the reach of humanitarian agencies.

Brief History

As with most conflicts in Africa, the current situation is strongly linked to the legacy of colonialism. The conflict's history starts with the violent 1885 Belgian imposition of colonial rule under King Leopold II, whom himself never visited the region.

After 75 years of colonial rule, the Belgians left very abruptly, relinquishing the political rights to the people of Congo in 1960. However, independence did not mean that economic rights enabled all inhabitants to benefit from the rich resource base. Still, former white colonialists dominated the economic, political and especially military ordering of the country. Furthermore, conflicts between different ethnic groups and regional powers about the new governmental structure arose. A few months after Patrice Lumumba, head of MNC, Congolese national libertarian party, became elected head of state, he was overthrown with US and European support by his former ally, Mobutu Sese Soko. Besides his claims for more independence for the Congolese people and his accusation of the former colonialists, Lumumba was suspected of cooperating with the Soviet Union during the Cold War period.

Mobutu used his U.S.-supplied arsenal to repress his own people and plunder his nation's economy for three decades, until his dictatorship was overthrown by the AFDL (Alliance des Forces Democratiques pour la Liberation du Congo-Zaire) led by Laurent Desire Kabila with the aid of Rwanda, Uganda, Angola, Burundi and Eritrea in May 1997. Kabila, also backed by the US, was accused by Congolese soldiers, Congolese Tutsi Banyamulenge¹⁴, Rwandan, Ugandan and some Burundian government troops of turning into a dictator of mismanagement, corruption and supporting various paramilitary groups who oppose his former allies. Therefore these groups themselves formed various Rebel Groups, opposing the government on the internal level since August 1998. As the conflict had raged on, rebels controlled about a third of the entire country (the eastern parts).

¹⁴ Banyamulenge is how the ruandish-speaking inhabitants of South-Kivu call themselves since 1967 to distinguish themselves from hutu-refugees of Ruanda-Urundi, now settled there, too. Literally translated it means "inhabitants of Mulenge."

Until the assassination of Laurent Kabila in January 2001, Angola, Zimbabwe, and Namibia supported the Congolese government, while the rebels were backed by the governments of Uganda, Rwanda, and Burundi.

Various African states, primarily South Africa, tried to intervene. In the Organisation for African Unity (OAU, today African Union) the "Conflict of the Great Lakes" has always received special attention. All these efforts lead to the signing of the Lusaka Ceasefire Agreement in 1999, which is considered to be the base of all peaceful solution proceedings in the conflict. Nevertheless, combats did not stop and peace was fragile. There were various political problems in trying to get a UN peacekeeping force in to help out, while killings continued. The UN deployed a small cease-fire monitoring body called Interim Emergency Multinational Force (IEMF) in 1999 which was upgraded to the UN-Mission MONUC (Mission de l'Organisation des Nations Unies en République démocratique du Congo) in July 2003 to 'protect civilians under imminent threat of physical violence'.

However, Amnesty International, amongst others, has noted that "MONUC has been a hostage to its weak mandate and has lacked the necessary equipment, personnel and international political backing."¹⁵

On January 16, 2001 Laurent Kabila himself was assassinated and his son Joseph Kabila became the new President of the DRC. He said that he would further encourage the need for cooperation with the United Nations in deployment of troops, strengthen the dialog of national reconciliation and help revive the stalled Lusaka peace agreements.¹⁶

The so called "*Innercongolese dialogue*", held from February to April 2002 in Sun City, was supposed to comprise five components, two rebel movements (the Uganda-backed MLC as well as the Rwandan-backed Congolese Rally for Democracy), non-armed opposition groups, political parties, civil society organizations and the government. But the power sharing question was mainly negotiated between the government, the rebel groups and one opposition group. The Lusaka agreements (the Ceasefire agreement of 1999 and its reformulation in 2001) were declared dead, as various groups had had disagreements on a variety of issues. But also the Innercongolese-Dialog in Sun City 2002 could not solve all of

¹⁵ <http://www.amnesty.org.uk/action/drc/international.shtml>

them, not to mention the struggles in the implementation of the arrangement. Several groups counteracted the implementation, others did not fulfill what they declared. Parts of the society who felt underrepresented funded new groups or parties, putting the whole process in question.

Nevertheless, the DRC is actually in the status of "Transition", with enormous aims:

First of all, a Reunification, pacification and reconstruction of the country has to take place. Once the armed conflict is stopped or at least limited and the direct war damages are reconstructed, the establishment of territorial integrity and the authority of the state over the whole territory are what is strived for. Of course, the relationship between ceasefire consolidation and controlling processes is debatable, but at least this is how plans are.

After the government and a territorial integration have been established, it is important for future coexistence that a national conciliation takes place. There, crimes of war will be brought to court and history is worked up.

The important role military played during every part of the conflict, especially thinking back to the overthrow of the democratically elected first national government of the Congo, the creation of an integrated and restructured national army seems to be a crucial point for the chances of success of the peace process. Its actual importance can also be observed, looking at the numbers of arms sold between all groups of society and foreigners. People are organized in Rebel-Groups and their subordinates or use light arms in daily life, in civil war.

For political purposes the organization of transparent and free elections for all levels, leading to the construction of a constitutional-democratic regime, is the predominant aim, but the process is hindered by various disagreements between society-groups and is nearly made into ridicule by the ongoing fighting, primarily in the eastern Kivu-provinces. At the moment, the date for the election is postponed from June 2005 to June 2006.

In general, the process is strongly restrained by the ongoing combats and disaccords concerning a diversity of economic and political interests between the following operating actors:

¹⁶ there has been a second Lusaka Ceasefire Agreement in 2001, which had also problems in its implementation.

- Joseph Kabila and Jean Pierre Bemba (MLC, noted on top) discussing the definitive arrangements in the state-structure with the other government parties
- RCD-Goma and UDPS (Union pour la Démocratie et le Progrès social) and especially the new "Alliance for the salvation of the Inner Congolese Dialogue" fighting for a reopening of the Dialogue to let all groups participate
- Meanwhile, UDPS and PALU (Parti Lumumbiste Unifié), the two greatest political parties in the DRC, prepare the organization of the political opposition, as they were too divided to find a common candidate for the elections
- Illegal exploitation of the resources in Eastern Congo is an important grade based on the civil war circumstances and the violent occupation of mines and other strategic territories by local groups
- The DRC inhabits more or less 250 different ethnic groups. The highest concentrations are Luba (18%), Mongo (17%), Kongo (15%) and Asande (10%). Ethnic questions are often instrumentalized for economic conflicts. The most critical case is the status of the Banyamulenge, Congolese soldiers of Rwandan origin.

Literature:

Aust, Björn: Feindliche Übernahmen. Ökonomische Interessen und "militärisches Unternehmertum" im Kongo, in: Azzellini, Dario/ Kanzleiter, Boris (Hrsg.): Das Unternehmen Krieg. Paramilitärs, Warlords und Privatarmeen als Akteure der Neuen Kriegsordnung. Berlin u.a. 2003, S. 143-160

Hofmeier, Rolf/ **Mehler**, Andreas: Afrika Jahrbuch 2000. Politik, Wirtschaft und Gesellschaft in Afrika südlich der Sahara. Institut für Afrika-Kunde, Opladen 2001

Scherrer, Christian P.: Genocide and Crisis in Central Africa. Conflict Roots, Mass Violence and Regional War, Westport 2002

<http://www.globalissues.org/Geopolitics/Africa/DRC.asp#BriefBackground>

<http://www.kongo-kinshasa.de>

<http://www.monuc.org/HistoriqueAn.aspx>

<http://www.globalissues.org/Geopolitics/Africa/DRC.asp>

<http://www.amnesty.org.uk/action/drc/international.shtml>

IS ANYBODY STILL A STATE IN CONGO?

REALISM APPROACH

Theories of *international* relations, such as Realism and Neorealism, are based on the interaction of states. Nevertheless, there are exceptions to these theories, which take “failed states” into account as well - where it is not yet clear to which state or power a territory belongs. In these cases, other states compete to fill the political vacuum with influence on the territory. This seems to be a possible scenario for the Congo Conflict. Hence we discussed in our workgroup the question, whether the Democratic Republic of Congo qualifies as a state under the criteria put forward by neorealism. An answer to this question would allow us to decide if the Democratic Republic of Congo is an actor in neorealist terms and thus relevant for explaining the conflict.

The Democratic Republic of Congo (DRC) was granted membership status to the UNO in 1971. Although this form of international recognition is usually a clear indicator for statehood, things are not as clear in the case of the DRC. So the workgroup tried to clarify the status of Congo by applying the neorealist criteria for states¹⁷

Concerning the first condition, *Administration (not bargaining)*, it has to be assumed that the Congolese Government does not have the capacity nor the authority to administrate the whole territory. The DRC actually is a state „in Transition“, with its first democratic elections planned for June 2005. Especially in the eastern part of the DRC there is an ongoing civil war concerning local domains. Therefore, the frontiers in this region are more or less blurred by migration and commerce. A constant unregulated flow of refugees pass the borders every day and the state does not have the means to control them. Thus, the criterion of *Borders, indicating domestic / foreign spheres* seems just as inapplicable.

The feature of *territoriality* can be seen in the context of these border conditions. Several times in the run of the conflict, refugee flows headed in all directions - for the central part of the country as well as the neighbouring states. Bearing this back and

¹⁷ States are organizations characterized by certain attributes: administration (not bargaining), territoriality (not nomadism), internal sovereignty, esp. power monopoly, social homogenization (not patchworks etc.), external sovereignty (no interference in domestic affairs), citizenship (not multiple

forth movement in mind it is highly unlikely that all the local social groups, partly belonging to transborder ethnics, attribute the same importance to the frontiers as the international society.

This brings us to the issue of *national identity (not regional or other)*. As the conflict has lasted for decades now and the complexity of ethnic, regional and national relations has reached to an impenetrable level, many locals orientate themselves towards primordial identity of family ties, being the only stable point. Even worse are situations in which these last family ties are destroyed violently. The created vacuum is often times filled by integrating into a military group - as in the case of the infant soldiers in the *Maji Maji Militia*. This militia consists in great parts of children, who are orphans or have been separated from their families in refugee camps or rural settlements. In this setting, the characteristic of *citizenship (not multiple identities)* seems to be even more futile. For example the question about the citizenship of some of the collectives involved in the conflict has been brought up repeatedly: Were the combatants Congolese or were they of Rwandan or Ugandan origin? Which groups of refugees have the right to settle in the DRC, which ones have to be accepted by the neighbouring states? On the one hand, these questions stand for the national and international significance that is attributed to a Congolese citizenship. On the other hand, and this seems to be more relevant here, these questions point towards the fact that there is no unambiguous concept of a Congolese citizenship to refer to.¹⁸

By internal and external sovereignty, we mean an internal power monopoly by the government and its apparatus without interference in domestic affairs by foreign powers. Looking at internal sovereignty, the most striking feature seems to be the disputes within the transitional government, consisting of president Kabila and four vice-presidents representing different opposition groups. This integrated body has not led to the expected unification but has actually deepened the rift between the groups in control of the territories. Consequently it cannot be assumed that the government controls the DRC as a whole. Apparently the integration of the different regional oppositions at the federal level in Kinshasa only diminished the level of

identities), national identity (not regional or other), borders, indicating domestic/ foreign spheres, symbols.

¹⁸ Read for example Antoine Lawson: " Central Africa: Integration suffers setback" in Pan African News Agency, March 17, 2004.

informality and the violence apparent in some confrontations. Nonetheless, there still remain local fights.

For the second characteristic, external sovereignty, it is important to note that the rebel groups dominating a great part of the DRC have much closer ties with Uganda and Rwanda. These connections have a stronger influence on their behaviour than their governmental participation in the DRC. This prevents them from following national interests for the DRC. In addition to this, Rwanda and Uganda are also capable of influencing the proceedings in the DRC directly. Another clear sign for a lack of sovereignty is the installation of the UN mission MONUC in the DRC and its further remaining until today.

Concerning the criterion of *social homogenization (not patchworks etc.)* the situation seems to be closely related to the points of *identity, territoriality and citizenship*. The DRC is inhabited by a variety of different ethnics, who play an important role in the conflict, or are instrumentalised for economic or military aspects. Any way, one cannot speak of social homogenization, as the social differences among the population are tremendous.

Regarding all the trouble with the named criteria so far, the last point, *state symbols*, seems to be of little relevance. There are state symbols for the DRC (a flag, a hymn etc.), which have symbolic meaning for the Congolese Citizens, but as long as the state still is "in transition" and the conflict situation impedes social development, a deeper identification with the state might appear impossible.

Having studied all these attributes, a clear decision in favour of one of the two possibilities laid down by the hypothesis is difficult. Whether the DRC can be considered a state acting (sovereign) in the conflict or whether the DRC represents a "failed" state and the only decisive actors are the other states involved in the conflict is hard to tell. On the one hand, the DRC does not completely fulfill a single one of the criteria. Hence the first possibility can be excluded.

However, the role of the construction of a state or the "transitional state" of the DRC is of such great importance in the conflict that it cannot be neglected in its handling. The territory has a history of state-constructions since its (de)colonisation; it is registered in the UN states list and is recognized all over the world. More importantly,

its lack in sovereignty does not mean that it has no power at all. Rather the impact of its political and economical weight, its military staff and the limited social infrastructure on the course of the conflict is apparent.

Accordingly the latter alternative, a "failed" state in hands of others, is also inapplicable. Negotiations about the terrain or the economical, social and political circumstances therein, without a decisive participation by the DRC itself seem to be unthinkable.

Ultimately, the case of the DRC status in the "Conflict of the Great Lakes" seems to fit none of the actor categories proposed in the Neorealism hypothesis discussed here. It exemplifies the amalgamation of the levels of analysis and the demand for more complex investigation frames adaptable to current conflicts.

APPLICATION OF PUTNAM'S "TWO-LEVEL-GAME" THEORY TO THE CONFLICT IN DRC

The Conflict in the Democratic Republic of Congo (DRC), or the "Conflicts of the Great Lakes," as it is often called in international press, seems to have been quite immune to any attempts towards finding an international solution. Particularly, the two most important ceasefire agreements, contracted in Lusaka 1999 (LCA99) and Sun City 2002 (SC02), failed in their application, although, when signed, aroused a lot of hope.¹⁹ In order to understand the dynamics of those negotiations and the reasons for their failure, we will now look at them applying Putnam's Two Level Game. We will start with the LCA99, as this one is more easily applicable because there is an international and a distinct domestic level involved. The SC02 on the other hand, called "Inner-Congolese Dialogue", did not really have an international level, since the negotiations mainly took place between different domestic opposition groups and the central government. We will return to that later, when looking for an alternative for the failed LCA99, which we are analyzing first.

The governments who signed the LCA99-agreement (the Democratic Republic of Congo (DRC), Rwanda, Uganda, Angola, Namibia, and Zimbabwe; see also the introduction to this conflict) constituted the first level, taking place on the international

¹⁹The German Office of Foreign Affairs still considers the LCA of 1999 as the base for all ongoing peace efforts in the DRC http://www.auswaertiges-amt.de/www/de/laenderinfos/laender/laenderausgabe.html?type_id=11&land_id=85, Downloaded on 30.08.2004

sphere. At the second (or domestic) level, where the "ratification" process took place, the rebel groups in the DRC, such as RCD and MLC, can be regarded as the most important actors besides the government and the political opposition. The demands of the rebel groups, which were or could not (be) met in the agreement, seem to have made the implementation of the planned peaceful arrangement impossible. Instead, one might get the impression that the rebel groups preferred the violent status quo, which according to Putnam's theory would then be a "voluntary defection".²⁰

Thus, our hypothesis resulting from a Two Level Game perspective would be the following: *As the Rebel Groups in the DRC opted for an ongoing war, the LCA 99 was spoiled in the ratification process by their "voluntary defection".*

How can we prove our hypothesis? First of all, it has to be noted that the natural resources within the territory of the DRC and their illegal exploitation are of great importance in this war. This has been discussed in detail in the „*Final report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo*“ by the UN²¹. Gertrud Kanu and Iseewanga Indongo-Imbanda quote the principal conclusion of this report: *"the conflict in the Congo centres mainly on the access, the control and the commerce with five important resources: Coltan, diamonds, copper, cobalt and gold"*.²²

Second, the close relation of the two levels makes a clear distinction between the political spheres more complicated and gives those actors advantages in following their interests, which are more or less directly present at both levels. We refer to the linkages between Rwanda and the RCD or Uganda and the MLC.

There is evidence that the RCD and the MLC were forced by Rwanda and Uganda to sign the LCA99. This made it possible for Rwanda and Uganda to uphold a responsible and peaceful image on the international level, being important for their international reputation in commerce and politics. On the other side, the implementation of the LCA99 failed exactly because the rebel groups refused to accept the consequences. They continued with the illegal exploitation of resources in the terrains they controlled and almost to the same degree still control today. As

²⁰ http://www.kongo-kinshasa.de/geschichte/ge_05.htm, Downloaded on 11.05.2004

²¹ Resume of the Final report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of the Democratic Republic of the Congo, 2002: http://www.welthungerhilfe.de/WHHDE/wir/positionen/Kongo_B_rgerkrieg_als_Wirtschaftsfaktor.pdf, Downloaded on 2004/08/30, further explications by Frank Nyakiru, in The Monitor (Uganda) of March 22, 2004

²² www.kongo-kinshasa.de (downloaded on 2005/03/03), translation by C.S.

observed by Gertrud Kanu and Iseewanga Indongo-Imbanda, *"the armed opposition departed several times from the agreement of Lusaka, by among other impediments continuing with the combats, leading to the conquest or the occupation of territories controlled by the Kinshasa-government"*.²³

All this was facilitated by the weakness of the Congolese state, which is financially supporting some of the rebel groups but still cannot control neither all of them nor its own territory. It is obvious that the state in general lacks power, it has no central control over the different regions officially subordinated to either the government or a rebel group. This situation results in many local fights.²⁴

Gertrud Kanu and Iseewanga Indongo-Imbanda describe how president L.-D. Kabila did neither act in favor of the LCA99-application: *"Before the assassination of President L.-D. Kabila on 01/16/01, the conflict parties departed from the fragile peace-agreement frequently. President L.-D. Kabila for example refused to accept the "Facilitateur" of the Inner Congolese dialogue, Ket Masire, for the reason that he regarded him to be one-sided and demanded a revision of the Peace agreement from Lusaka immediately. He also offended the positioning of the UN-mission in the DRC (MONUC)."25*

The instability of the government and, first of all, the resistance of the rebel groups against governmental sovereignty, are not determined by political purposes that could be handled on the first level. What is of interest for the parties in the conflict is rather the local control of resources and commercial routes, which determine the military (and political) strategy of the actors. If these interests are more easily reached in a state of civil war, then every attempt for peace will be counteracted by a "voluntary defection."²⁶

Due to the fact that economical circumstances in wars transform the interests of actors, the problem can be that those circumstances are not approached with peace attempts referring to the principal origins of conflict. The growing importance of the "economy of war" has come to overshadow the principal objectives.²⁷

²³ www.Kongo-Kinshasa.de (downloaded on 2005/02/14), translation by C.S.

²⁴ Vgl. Aust, Björn: Feindliche Übernahmen. Ökonomische Interessen und "militärisches Unternehmertum" im Kongo, S. 145

²⁵ www.Kongo-Kinshasa.de (downloaded on 2005/02/14), translation by C.S.

²⁶ Vgl. Aust, Björn: Feindliche Übernahmen. Ökonomische Interessen und "militärisches Unternehmertum" im Kongo, S. 149

²⁷ Vgl. Aust, Björn: Feindliche Übernahmen. Ökonomische Interessen und "militärisches Unternehmertum" im Kongo, S. 149 and Vgl. Ballentine, Karen: Beyond Greed and Grievance:

Nevertheless, the rebel groups can not be seen in total confrontation to the government of Kabila. Kabila's demission was utopian and in addition to that would also have provoked a very risky situation for the rebel groups themselves. The abrogation of all legal frameworks would have led to even more intense fighting among those groups and in consequence could threaten their assets.

Which aspects would then be of interest to the rebel groups, permitting the government to broaden its win set?

An autonomous status of the territories they already control in Eastern Congo (esp. Ituri and North- and Southkivu) would legitimate their predominance in these regions and would give them more economic and political liberties, especially in confrontation with the UN-troops. An integration of the most powerful rebel groups into the government would certainly lead to the manifestation or even extension of their influence.

These were exactly the issues discussed in Sun City: the governmental spheres were divided among the different parties of the "Inner Congolese dialog". The government now consists of President Joseph Kabila (son of Laurent-D.), his staff and four vice-presidents representing different opposition and rebel groups. This structure, however, has not led to the expected unification between the different groups towards a coordinated coercion of power, but has actually deepened the division of the territory between the participants. So the official government still cannot claim control over the whole DRC, although the integration of the different regional opposition groups at the central level in Kinshasa decreased the informal space where many groups acted and therefore diminished the extent of violence in some areas. Nevertheless, there remain local fights at an alarming intensity.

With that concept, SC02 made progress in the conflict and generated hope towards a real peace process. But again, its ratification on the internal level had to pass unexpected difficulties. Relatively new actor-alliances, such as the "*alliance for the salvation of the inner Congolese dialog*" and "*the Congolese Opposition*", fought for an inclusion of all parties into a new-opened dialog. Besides those groups gaining importance, controversies among the already integrated groups continued.²⁸

Reconsidering the Economic Dynamics of Armed Conflict, in: Ballentine, Karen/ Sherman, Jake (Hrsg.): The Political Economy of Armed Conflict. Beyond Greed and Grievance

²⁸ Controversies between J. Kabila and Jean Pierre Bemba (MLC) about the formalisation of the new constellation: <http://www.kongo-kinshasa.de/geschichte/geschichte5.php>, downloaded on 2005/03/10.

This again shows the complexity of the conflict in the DRC and the importance of a cautious and detailed consideration of every local fight and economic circumstances.

All in all we get to the conclusion, that the hypothesis seems true and the Two-Level-Game-theory is useful to explain the ongoing of the Congo Conflict. Central for us was the concept of voluntary defection. Economic interests and the longing for autonomy of the decisive domestic veto-players have spoiled the agreement reached at the international level. In the Application of Putnam's Two-Level-Games-Theory, it also becomes very apparent how the complex political and social constellation of the DRC is determined by the tremendous illegal exploitation of natural resources. Any attempts for a peaceful solution have to take this into account.

One could also think about concentrating efforts on the reduction of the illegal exploitation of resources in the first place, which would be the only other possible solution of the situation. This would also require a more intense dealing with the illegal negotiations between the conflict groups and their partners in Rwanda, Uganda and in the "industrialized" countries.

Confronted with the weakness of the Congolese state, one could, in this context, even doubt the possibility of the distinction between the two levels that are central for Putnam. However, for us it seemed reasonable to apply the Two Level Game as the Congolese state has reached internal and international recognition and is actually far from being totally deconstructed.²⁹

²⁹ Although one should not forget that its construction is mainly based on colonial history and international, especially UN, construction.

THE CRISIS IN COLOMBIA

BY ONDŘEJ SPAČEK:

COLUMBIA – AN INTRODUCTION



Source: Central Intelligence Agency,
http://www.lib.utexas.edu/maps/americas/colombia_pol_2001.jpg

The members of the Latin America Group decided to focus on the crisis in Colombia. We chose this conflict area in spite of the fact that the conflict we are witnessing here is not a purely international one.

Colombia is located in the north-west of South America bordering Panama in the north, Ecuador and Peru in the south, Brazil in the south-east and Venezuela in the north-east and has a direct access to both the Pacific and the Atlantic Ocean. In large parts of the

state area, its inhabitants face difficult living conditions, especially in the Andes mountain range in the western part of the country, which makes up about one third of the total area, but where about 80% of the inhabitants live. Poverty is also widespread in the swamps of Amazonas in the south-east covering another third of the country's territory. The Andes are divided into three subsystems by the rivers of Magdalena and Cauca: the Cordillera Colombiana Occidental, Central and Oriental. Other important rivers are the Orinoco, constituting part of the border to Venezuela, and the Guaviare in the south at the entrance to the Amazonas region.

Despite its extensive area (1 141 748 km²) and a very uneven population allocation, Colombia is a unitary state with 32 departments administered from the capital district of Bogotá, located in the mountain region in the centre of the country. However, large parts of the country are not controlled by the central government, but are occupied by paramilitary forces such as the AUC (*Autodefensas Unidas de Colombia*) or by

guerrilla groups, the biggest of them being the FARC (*Fuerzas armadas revoucionarias de Colombia*).

Similarly to other countries in the region, national politics have been characterized by the conflict between liberals and conservatives right from the foundation of these parties in Colombia in 1849, including some armed encounters (e. g. the “The War of the Thousand Days” between 1899 and 1903, ending with Panama (“Nueva Granada”) being split up and therefore setting a definite end to what once used to be “Greater Colombia”). This bipolar structure persisted well into the 20th century reaching its climax in the conflict known as *La Violencia* between 1949 and 1958, leaving behind 250 - 300,000 dead. The two parties agreed to cooperate in a National Front where the presidency rotates and cabinet seats are divided equally. This agreement lasted formally until 1978 and practically until 1986, when new actors appeared on the Colombian political scene.

Some analysts argue that it was explicitly the agreement between the two most powerful parties that contributed to the formation of armed opposition. In 1965, the ELN (*Ejército de Liberación Nacional*) was founded, as the first guerrilla organization in Colombia. It was followed by the FARC (*Fuerzas Armadas Revolucionarias de Colombia*) in 1966, the military wing of the Communist Party of Colombia. Today, it has some 17,000 members, being the guerrilla movement with the biggest political influence. Finally, the M-19 (*Movimiento 19 de Abril*) was founded in 1971, the only guerrilla that managed to transform itself into an ordinary party by an agreement with the government in 1989.

The guerrillas, politically far left, have to finance their activities. Drug trafficking and later kidnapping became the main source of income, although the majority of guerrilla leaders still deny receiving any financial resources through drug business. What started as political opposition has developed into a conflict touching the whole society and every sphere of life.

Throughout the years, a number of attempts to stop the conflict, by some even called civil war, were undertaken. In 1982, for instance, President Belisario Betancúr Cuartos granted amnesty for the guerrilla combatants and freed political prisoners. Neither this, nor a Peace Commission in the 1990s was very successful. The successful institutionalisation of M-19 was only of temporary importance as the party didn't manage to keep its active role in the political life and was last represented in the parliament in 1991.

A new series of peace talks started in 1999 after the Conservative Andrés Pastrana Arango had become president in the 1998 elections. The FARC was granted a safe zone of the size of Switzerland. To negotiate a lasting peace agreement, Pastrana and the FARC leaders met in 1999. In October 2001, the representatives of the Colombian government and the FARC signed the San Francisco agreement. The parties committed themselves to negotiate a cease-fire, but the negotiations soon had to be interrupted due to the lack of compromises between Pastrana and the FARC and the constant violations of the treaty. Pastrana's "peace experiment" ended with a military invasion of the formerly autonomous territory of the FARC and numerous casualties on both sides.

From what has been said until now, it could be assumed that the Colombian conflict is rather an internal problem. Our working group, too, was tempted to concentrate on the conflict in Colombia itself. However, in the end, we managed to see the international side of the conflict.

The US-American Administration and especially the US Drug Enforcement Agency turned out to be the strongest player on the international level (3rd level of analysis according to Waltz). The United States of America have been active in the region since the 1960s, when they started granting financial and military support to South American governments (and later also to rightist opposition groups, as for example in Chile in early 1970s). The US employed these means in order to limit the spread of what was perceived as a Cuban-inspired revolutionary threat and to support the "democratic candidates", often ignoring the difference between a leftist democratic party and armed groups.

In Colombia, US military advisors were quite active for the government of the country and inspired the formation of paramilitary forces of "self-defense". In 1965, a presidential decree legalized the armament of civilians. Three years later, this decree was converted into permanent legislation and the law stayed in force until May 1989 when it was declared unconstitutional by the Supreme Court. Later on, penalties for members and organizers of paramilitary groups were introduced in an additional decree.

After the end of the Cold War, the priorities of the United States in South America, and especially in the Andes, changed and the main task became to eradicate the cultivation of coca, its processing into cocaine or heroin and the subsequent distribution on the US market. Colombia, today the world's largest cocaine producer

and supplier of about 90% of the cocaine and up to 60% of the heroine distributed in the United States, was a clear target.

Since the beginning of the 90s, huge sums of money have been invested and different approaches (aerial spraying of the coca fields, incorporation of US forces in the Colombian army, several talks of the Colombian government with the guerrilleros) have been applied to stop the drug trade and normalize the situation in the country. Until now, with very limited success. On the contrary, the US policies are provoking a lot of criticism especially concerning the militarization and criminalization of the Colombian civil society through US-politics and policies.

CONSTRUCTIVISM – THE ALL EXPLAINING WONDER THEORY?

In the session on realism, the group unanimously agreed that the (neo-)realist black-box vision of the world does not give enough clues for explaining the Colombian conflict. As an alternative meta-theory, constructivist approaches are challenging realism. Constructivism does not solely open up the black box of the state. It also comes with a new perspective on what is happening in it. Constructivist authors concentrate on the social structures and on the way how they are being created by the actors and how they influence the actors in return. Not only material power of the actors is taken into account. The constructivist researchers concentrate also on the “discursive power”, power of knowledge, ideas, culture, ideologies and language. Constructivists analyze the power defining identity from which then the interests of actors are derived and which inspires the form of social structures. All in all, we were confronted with a completely different approach to international relations and were eager to see if it could deal more successfully with our conflict.

In our view, it surely could do this in one aspect. It helps us to understand the relationship between Colombians and their own country, the neighboring states and the United States of America. From our point of view, this relationship is based on the idea of nationalism, which can be only taken into account when using a constructivist approach in the analysis. Let us return to the time of the liberation fights in the 19th century and have a more detailed look on this phenomenon. The struggle for independence was started by the Creole³⁰ elites of the colonies and their motivation was clear: better career chances for themselves and their children, because under

the Spanish (or Portuguese) rule, the future prospects of Creoles were much more limited than those of peninsular Spaniards. Whereas the *peninsulares* could assume higher positions in the civil service both in Spain and all the colonies, the Creoles were limited to their own vice-royalty. The liberation of Latin America was elite-led and the pan-American coalition among the newly independent nations split up shortly after the victory was gained.

Afterwards, each country went its own way in creating national heroes, symbols and legends in order to form a national identity. In Colombia, for example, the national hero of the first order became Simón Bolívar, the main protagonist of the liberation of the vice-royalty Gran Colombia. The strategy was successful – all the countries were able to construct relatively strong national identities and create patriotic feelings inside their citizens. This sometimes led to animosities and war, as for example between Argentina and Chile or Chile and Bolivia. However, the feeling of common past prevailed and it was reinforced by the image of a common enemy. First, it was Spain, but with its fading power, it was replaced by a new world power – the United States of America.

The constructivism also makes it possible to have a closer look at the ideologies of the actors in the conflict. The degree of influence of the communist ideology on the guerrillas can be analyzed, but in our point of view is rather low. Also, the importance of ethnic origin for taking part in the conflict can be questioned in a constructivist approach. We believe that ethnic origin does play a role, as for example the majority of the members of the Colombian government is white. This aspect isn't decisive, however. A constructivist research might even bring ideas how to solve the conflict, for example by reaching a "national reconciliation" dealing with the values and symbols that are common to all the actors of the conflict.

An important disadvantage of the constructivist approach is the high proportion of field work that has to be done. This is always connected with high costs on time and other resources, in the case of Colombia additionally with substantial security risks. A simple analysis of the discourse using newspaper articles, speeches of the main personalities of the parties participating in the conflict or of the statements for the press would probably not offer a lot of new information about the present status of

³⁰ A Creole is a white person born in the New World.

the conflict. However, it could well be used for an analysis of the history of the conflict. We would be able to see its evolution and maybe even find some regularities in time, phases of lower and higher tension between the actors, and thus predict how the conflict could evolve in the future and say when chances to solve the conflict are highest.

Constructivism might seem as an all-explaining wonder theory on the first sight, but its complexity and its need for lots of primary data makes it, at least in our eyes, applicable only to partial aspects of the conflict and not to the conflict as a whole.

GLOBALIZATION

The lecture on globalization proved to contain some very applicable aspects to our conflict. The “*process generated by world wide interplay of capital flows and communications flows enabled by new technologies*” can be noticed clearly not only in the big industrialized countries of the Northern hemisphere, but as well in a small South American country seemingly caught up only in its own national business.

Primarily, the group discussed the direct influences of globalization on our region by looking at the relevant flows. Concerning the first two examples mentioned in the definition, capital and communications flows, we found out that big loans by the IMF and the World Bank constitute a great part of the international *capital* flowing into Colombia. Secondly, the money obtained from the international smuggling and trafficking of drugs and weapons makes up another substantial part of the Colombian shadow economy. Certainly these capital flows are a consequence of globally formed networks and interests.

As another example, the *human* flow of migrants out of Colombia is quite astonishing: In 2002 alone, more than 360 000 Colombians applied for entry into the USA. The rising migration numbers are alertly observed by Amnesty International and other NGOs, which form another network trying to regulate *communication* flows out of the country. Together with the UN and representatives of the EU and US, they keenly observe the implementation of human rights in the country and put pressure on the government to apply democratic practices and reliable jurisdiction.

The other question we discussed focused on the existence of realistic exit options for the region and its actors to avoid being dragged into an unwanted process of globalization and form autonomous regions free from international pressure. To analyze this question we mainly focused on the economic side of globalization: A common South American market like Mercosur would enable the participating countries to protect themselves from the US domination of prices. Apart from that, we found inner-state peace and security to be a prerequisite for further sovereignty of the state. Only if Colombia manages to find a solution for the ongoing conflict, will the country have a right to refuse the international interference in their issues. In this situation direct international influence might be at least reduced, but the question remains, whether it could prevent the country from becoming subject to the various effects of globalization. Having thus argued, the group came up with the question of whether our conflict might somehow even be promoted by globalization and the international interest in the country, or whether these factors help to constrain the conflict.

Finally, we spoke about the relations of international dependence and recognized our country as being clearly on the dependent side of globalization. As many of the Latin American states, Colombia exports mainly agricultural products like coffee or flowers and depends on the international trade system to obtain rather expensive imports in the field of technology, which are not being substituted by national industries. Furthermore, the lacking personal security, bad infrastructure and absence of skilled workers do not attract international investors and force many people out of the country, which eventually results in the high amount of migration mentioned above.

Summing up the arguments, the group came to the conclusion that Colombia, like many other developing countries, does not shape the process of globalization, but instead is constantly shaped by it. While we detected capital flows, content flows and information flows into Colombia, the only constant flows out of the country unfortunately seemed to be drugs and migrants. Since the territory holds no important hubs or nodes for financial or informational conjunction, it cannot be expected to find a way out of this position in the near future, but will have to fiercely work on solving its inner conflict in order to profit from the process of globalization.

The Significance and Roles of Teaching Theory in International Relations



Introduction

With the increasing independence of the field of study called International Relations (IR) and the professionalisation of its teaching, theory courses have acquired a more prominent place in the newer curricula. Some observers have started to question whether this can be justified. Criticism against an overly theoretical nature of the studies typically revolves around one central theme. Theory is said to be most divorced from practice. Hence, for teaching to be useful in the real world of the future practitioner, it should recoil from (pure) theory courses and emphasise applied studies (Wallace 1996).

This article tries to show that such an argument is based on an erroneous understanding of the relationship between theory and practice, and of the specific roles theory can play in the education of practitioners. For it conflates the theory/practice distinction with the academia/politics distinction. Hence, theory seems to be coupled with the “ivory tower” of academia, whereas applied studies are linked to (diplomatic) practice. Although some theoretical work will undeniably be scholastic, and some empirical work of

practical use, this is not necessarily so. A quick look at much of the empirical literature in IR in, for instance, *International Studies Quarterly* or the *Journal of Conflict Resolution*, bears witness to the remoteness of applied studies from any direct practical value. On the other hand, theoretical research is of great political significance and commands public interest as, for example, democratic theory or theories of globalisation.

This article will reveal the neglected roles that the teaching of theory in IR can fulfil. These roles support the claim that there is a strong need for teaching (and researching) theory in IR. Indeed, this article will argue that thinking in terms of these very oppositions between theory/practice and academia/politics is part of the problem for understanding the role of academic education in IR today. For they lure us into the illusion that empirical studies can ever be divorced from theory, and that university education is only there to serve the policy relevance of the day and not the intellectual breadth and maturity of future practitioners.

My first argument is logical: no empirical analysis is without theoretical assumptions, without a framework of analysis. There is no explanation that is simply a neutral selection of “data speaking to

us". Related to this is the dual character of theoretical knowledge: it is both explanatory and constitutive (Smith 1995:27-8). In its classical explanatory sense, social science theories are the result of knowledge giving a common, more general and coherent explanation for a variety of specified cases. But this does not exhaust the function of theories. Theories also have a constitutive function; *i.e.* a theory is the condition for the very possibility of knowledge. Without concepts that cut through the forest of empirical data, we would be unable to see the wood for the trees. Theories are not just the result but also the precondition for the possibility of empirical knowledge.

My second argument is educational: making future practitioners and observers aware of the constitutive function of theories fulfils the crucial role of a more time-independent intellectual education. As mentioned by Wallace (1996:317), the diplomat in the United Kingdom used to be trained through classical studies that gave them the general and time-independent skills to decipher and respond to changing political situations. The point here is simply that today we need to update this approach. Besides the necessary factual training in international law, history, economics, and politics, future observers and practitioners in international affairs (who might not necessarily be public servants) need to acquire the skill of intellectual self-distance, reflexivity as it were, to respond to changing challenges. Moreover, this ability and the related capacity to reflect on one's own and another's assumptions are crucial for the tasks of understanding and negotiating across national boundaries.

On the basis of these two arguments, I will consider the role that teaching and research in theory can play for entire academic communities. The article will try to show that the neglect of theoretical

studies can cement the peripheral position in which the majority of academic communities find themselves today. If not opposed, the international division of academic labour tends to slot them into mere data providers and thought-takers. Theoretical expertise is, as some cases show, part of the way out.

Finally, I sketch out some of the implications this understanding of the roles of theory in teaching IR has on the type of theory teaching. I identify and discuss a non-exhaustive list of four types of courses where the constitutive function of theories is fruitfully addressed.

A Logical Argument: The Constitutive Nature of Theory and the Necessarily Theoretical Nature of All Knowledge

There is sometimes an assumption that 'theory' is something that is suitable only for 'advanced' students. The fear is that students are not interested in theory, that they study IR with a practical orientation and become alienated if asked to think conceptually and abstractly, and, most damagingly, that students want to be told the 'right' answers and not to be exposed to the scandalous fact that authorities differ even on quite basic issues. These positions must be resisted. All understandings of IR and of the other social sciences are necessarily theoretical, the only issue is whether this is made explicit or not and most good students are well aware that this is so (Brown 1997:vii).

How do we know what we know? This seems an arcane question and yet it is the basic question for establishing the fundamental identity of an observer or a scientist. For it allows us to justify why we believe something to be true. Teaching, in turn, has to do with the communication

The
Significance
and Roles of
Teaching
Theory in
International
Relations

of this knowledge – both its content and its means of justification.

This position makes it very difficult to conceive of science and teaching in a simplistic, empiricist manner. By simplistic empiricism, I understand the position that “data speaks for itself”, that is, that we can neutrally access empirical data. There is hardly anybody who subscribes to such a position in the philosophy of science, positivists included. Any empirical explanation relies on *a priori* concepts. The question then becomes whether the choice of such concepts, albeit necessary, can be neutral or innocent with regard to the event to be explained.

Going after the business of empirical research, some scholars, however, tend to bracket those questions. In IR, possibly the most famous research programme in this more empiricist tradition is the Correlates of War Project. This project is led by David J. Singer, who has succeeded in obtaining an almost incredible amount of resources over the last few decades, seeks to find out which antecedent conditions correlate with war.¹ The project is based on a huge historical database of international conflicts for which we have enough information to code them. It is inductively driven in that it wants to derive knowledge from empirical correlations. In other words, our knowledge is based on empirical generalisations of which antecedent events correlate with war. In its self-understanding, this is the only possible way to get unbiased information.²

Apparently absent, theory enters twice into this type of explanation. First, as empiricists themselves stress, theory is needed since these correlations do not explain anything in the strict sense of the word. For they do not answer the question of why things correlate. Only an argument about causes can help us find out whether the correlations are spurious,

or whether they are the social science equivalent of laws. It might be added that even economists, admittedly institutional ones, are not that certain whether social sciences can actually find these more general theories. Richter (1994), one of the *doyens* of institutionalist economics, has likened economics not to physics, but to medicine where we still do not know the causes, say, of rheumatism, but have (*via* trial and error) discovered ways to mitigate its effects.

More importantly perhaps, theories enter the analysis already before or rather for the establishment of these correlations. As already mentioned, we need concepts to code these events. Without concepts as meaningful data-containers (Sartori 1970), we cannot distinguish music (a meaningful fact) from sheer noise (the totality of information) in world history. In other words, pure induction is not possible. In turn, such concepts simply cannot be divorced from theoretical or pre-theoretical assumptions. This is also called the necessary theory-dependence of facts. How do we know, for instance, that the things we compare over the millennia, and which we label with the same concept (in this case, war) are actually the same? Did they mean the same to the actors then and now? The very possibility of “conceptual stretching” (Sartori 1970; Collier and Mahon 1993) is dependent on certain assumptions about history and/or human nature, for instance.

This criticism of empiricism does not necessarily imply that reality can be reduced to what we think about it. In the natural world, reality itself does impose limits on the way we can understand it. In the social world, which is more of “our” making, not every explanation will reach at least some intersubjective consent as being plausible.³ Hence, if pure induction is impossible, if facts are always theory-

Stefano
Guzzini

dependent, this does not mean that the real world can be meaningfully described in a completely arbitrary way. The offshoot of the previous discussion is neither that there is only one true explanation for everything nor that "anything goes", but that there can be a series of plausible and theoretically founded explanations for which, at any given time, we might not have enough evidence to decide between. Indeed, to put it more strongly, it cannot be excluded that these explanations will approach the apparently same project from very different angles, asking often incompatible research questions about it. In this case, there would be no common evidence against which we could make a final assessment.

This should not be confused with a related argument that ideologies imbue empirical research and that ultimately no justified choice between such ideologies can be made. This reading of the theory-dependence of facts is a very lazy attempt to stop any scholarly communication between, instead of just among, true believers (Guzzini 1988). Such a reading has been reinforced by the classical American way of framing IR/International Political Economy theories according to the triad of political ideologies in the United States (US): conservatism, liberalism, and radicalism. But such a confusion of (a particular national) ideological debate with meta-theoretical assumptions is not warranted. There is, for instance, no reason to assume that conservatives will necessarily link up with realism. Keohane (1989) was perfectly right that one can be both a realist (in IR) and a liberal (in political terms). More importantly, it is a matter for debate to establish whether ideologies really distort empirical analysis beyond a common ground. So-called realists and idealists in the classical IR tradition shared many of their assumptions about the international

system; hence, they were able to analyse the state of affairs in a very similar way. They could differ about the question of whether or not this was a state of affairs to stay unchanged.

Instead, the importance of constitutive theorising and concept-formation better shows in the now inflamed discussion about what led to the end of the Cold War.⁴ The interpretation of this event can already start with the exact dating.⁵ Similarly to the debate on the origins of the Cold War, where scholars put the date at 1917 (Fontaine 1965) or 1945 or 1947, the end of the Cold War has been dated at 1985 (the rise of Gorbachev to power), 1987 (for some, the actual policy turn in the Union of Soviet Socialist Republic — USSR; e.g. McCwire 1991), or more commonly 1989 (the fall of the Berlin Wall, either symbolising the end of the Eastern bloc or the commitment to change), and 1991 (the end of the Soviet Union). As with the nature of the Cold War, the dating here reflects whether the Cold War is seen as clash of ideologies or of superpowers, and whether individuals or structures play a role in world politics. The very interpretation of what the Cold War was becomes an issue which cannot be taken for granted when asking the question.

The answers to the question are similarly imbued by theoretical assumptions. Those who see the end in the final demise of the Soviet challenge to the US supremacy will have a materialist understanding of power, most saliently the new round of the arms race launched by the first Reagan administration. The USSR, so the story goes, was forced to give in. Apparently, the pure material power figures are, however, too vague to allow such an interpretation, even for realists themselves.⁶ Therefore, some realists rescue themselves by saying that the USSR perceived a power decline and reacted

The
Significance
and Roles of
Teaching
Theory in
International
Relations

Stefano
Guzzini

accordingly, first stepping up efforts and then giving in (Wohlforth 1994-95), a claim now supported with a turn to political economy (Brooks and Wohlforth 2000-01). Still, this leaves a lot unexplained for understanding the extent of Soviet retrenchment (Kramer 2001), let alone about the origins of the Soviet legitimacy crisis in the 1980s. Another interpretation shared by those who conceive of the international system as a social construct in which governments behave on the basis of their self-identification will point to the non-material causes that pushed the USSR to change policy (Wendt 1992; Koslowski and Kratochwil 1994). In fact, the "New Thinking" opened many diplomatic avenues, earlier forestalled, simply by ignoring the in any case illusory military threat from the West. Still others will emphasise domestic politics, that is, the fact that it was only after the end of the Cold War, and not for any international anarchy or great power competition that the USSR disintegrated, but through the emerging Russian nationalism used by Yeltsin in his power competition with Gorbachev.

The assumptions which inform the interpretation are crucial for both the actual explanation of events and for policy advice. The interpretation of such crucial events constitute the "lessons" of history which then inform the judgement of many policy-makers. The example of the Great Depression might be useful here. For economic observers (Kindleberger 1973/1987), the Great Depression is mainly remembered because it showed the negative effect of closing markets off from international trade. Because countries did not work together to keep their markets open, their "beggar-thy-neighbour" policies only pushed the problem over to the next country from which, after a cycle, it would inevitably return in

an exacerbated form. These lessons focus on the level of the international economy and on the efficiency of markets. More politically oriented observers will remember the Great Depression for the political turmoil which could only be stopped by inventive state strategies, such as the first Swedish social democratic experiments and the "New Deal" (Strange 1998). If not countered, countries were ripe to fall into the hands of populist regimes, some of the worst sorts. Whereas the first vision, correctly, claims that state intervention in international markets deepened the depression, the second vision, equally correctly, would argue that state intervention was key in avoiding a turn towards authoritarian regimes in some countries (which would actually have been even more interventionist). These lessons still inform some of the debate about globalisation today. Defenders like Gilpin (Gilpin, with Gilpin 2000) point to the risks of protectionism (as if globalisation were mainly about trade). Even moderate critiques like Strange (1986; 1998), however, point to the necessity to "cool the casino", to manage "mad money".

Indeed, these ideas might bring about the very things they wish to portray or avoid. For a long time, peace researchers have been arguing that realism was a self-fulfilling prophecy.⁷ If all other governments assumed that Germany was prone to return to a more irredentist policy after its reunification, an argument forcefully defended by Mearsheimer (1990), then their policies might have isolated Germany to such an extent as to provoke a more assertive and aggressive Germany.⁸ If all governments assumed that the next big conflicts are "clashes of civilisation" (Huntington 1993), then their behaviour, opposing policies on the basis of a supposed threat to one's own culture, might well trigger ethnic/cultural conflicts which would not otherwise have

appeared — besides justifying “ethnic cleansing” as the only rational way to solve the multicultural “problem”.⁹

The Educational Argument: Teaching Future Practitioners

Exactly because data does not speak for itself, because all observation is theory-dependent, and because observation can in itself have an effect on this very reality it is supposed to describe, it is fundamental that observers of international relations, whether practical or academic, be trained to become aware of their own and others’ assumptions. As seen in the previous discussion on the end of the Cold War, these are highly significant questions for academia and politics. For this, they must understand both the explanatory and the constitutive functions of theories.

Nearly everybody agrees that teaching the explanatory or instrumental function of theory is important. For many scholars, this is basically the only path to theorising, its sole legitimisation. On the basis of case studies and comparative research, the discipline develops so-called middle-range theories. These are explanations for which one might be able to specify (scope) conditions under which they will with a certain probability apply to specified cases. They can also be used as tentative tools for policy advice. There, careful use of these models can produce a series of scenarios. It is through this ability of theories that science is usually considered “useful”.

But IR is no different to any other social science: these explanations are often limited. Not only do we lack a general theory of international relations or foreign policy: we can give no “eternal

advice” to the prince, as Aron (1962) already admonished. But given the amount of variables, and here the social sciences differ from the natural ones, we either tend to make reductionist explanations by concentrating on too few, or must refer to causal complexes without being able to disentangle the host of possible explanations. Forecasts are often nothing more than educated guesses. As Grosser (1972) once put it, political scientists are only able to tell afterwards why things had to happen that way (why a particular scenario was realised). Hence, if conceived similarly, and compared, to the natural sciences, IR theory frequently remains a very blunt instrument.

Nonetheless, there is another important utility in teaching theories in their explanatory function. It is a first step on the road to see the important constitutive function of theories. Insofar as this teaching emphasises the relationship between the event and its explanation, it helps to sharpen our mind on the way these very theories have been constructed. This is extremely important for learning to work theoretically, and to make this work more precise, a point on which again the majority of scholars would agree.¹⁰ In other words, such teaching is meant to develop the capacity of students to train in clear thinking. As a first step, students must be made aware of the very difficult, but crucial, step of concept formation: why do we take which concepts for which type of empirical analysis? What is the cost of choosing one concept over another? Then, students could be made aware of how such concepts are used for the explanation of events. That is, they could get a sense of how interpretivist historians work, how they construct and defend their claims. Finally, they could try out making an inverted research design of other works before trying it out for themselves.

The
Significance
and Roles of
Teaching
Theory in
International
Relations

Stefano
Guzzini

Looking at it in this way, we have already *de facto* reached the constitutive function of theories. This methodological understanding which stresses the problems of the crucial initial step of concept or ideal-type formation falls within the classical Weberian tradition.¹¹ Social science methodologies that skip this part, like King *et al.* (1994), tend to paper over the differences between naturalist and interpretivist approaches. By unproblematically using concepts, their positivistic research design takes for granted that which actually carries the most weight in interpretivist explanations.¹²

One of the obvious merits of teaching the constitutive functions of theory underpins the above-mentioned qualification and was already covered in the previous section. There, we saw that concepts and conceptual frameworks are "the condition for the possibility of knowledge" as Immanuel Kant put it. Moreover, we met the basic social constructivist tenet that ideas about social events and those very social events can interact. Of course, students must be made aware of this.

But there is a second, apparently more far-fetched argument in taking the teaching of constitutive theories very seriously: its practical use for international politics. Learning theories is a means for acquiring skills that are crucial for diplomacy. Let me support this claim by making a little detour.

Any knowledge at any given time is bound to be limited. Our knowledge today is prone to change over the next couple of years, and certainly over the next decades. Yet, students as future practitioners will have to be knowledgeable over time. Hence, they must be trained in their capacity to assimilate and produce knowledge on their own. They must be intellectually independent by the time they leave university. Acquiring intellec-

tual independence implies not merely practical knowledge about where to find information in the future, but the capacity to auto-correct one's knowledge. For this, being aware of one's assumptions, and how they relate to understanding, is absolutely fundamental. In other words, theoretical self-awareness is crucial for having the flexible and knowledgeable mind necessary to adapt to new and different practical circumstances. It is for this aptitude that traditional diplomacy often relied on classical education in the past and that many business firms in Western Europe have started to systematically recruit some of their staff from the so-called soft social sciences, and not only from law, economics or management. The same, of course, applies to public administration.

Yet there is an even closer link to diplomacy. Besides the necessary technical skills (law *etc.*) mainly necessary for the consular functions of foreign offices, other foreign office positions still have a touch of classical diplomacy. This applies in particular to the sections on information, some representational functions and, of course, all negotiations. There, diplomats are responsible for providing politically relevant interpretations about significant events in other countries, or for directly managing the relations between countries.

The classical education for these more diplomatic positions includes some training directly offered by the foreign offices, *i.e.* a form of in-house socialisation. Usually, accepted candidates undergo two years of courses and then qualify for their jobs. But such an approach presupposes the very existence of a foreign policy tradition. This is not an obvious option for new countries that often have to build up their foreign offices from scratch. Moreover, it assumes that socialisation is up to date with the way international diploma-

cy is run today or that it has strategies for correcting itself.

In countries with little in-house socialisation, universities take over part of this task. This happened in the US particularly after 1945, when the country's elite became even more self-conscious of its role in world politics, then during the days of the Princeton Professor turned President, Woodrow Wilson. The US academic solution after the World War II, as most forcefully represented by Morgenthau (1948), consisted of proposing a set of rules to follow in foreign policy-making which could be scientifically deduced from eternal laws (to be found in human nature). Science, not the clubby in-house socialisation was the answer. However, whereas in the past some might have believed that there is a science of foreign policy whose tools we simply have to apply, these certainties have withered. As a result, today we cannot simply propose a deductive science for the future diplomat, even if we have made progress in some domains.

Hence, some countries might find themselves facing the lack of both any history of their foreign office for socialising future diplomats, and of any deductive theory of diplomacy. So how do we train diplomats (not the consular clerks) and foreign policy specialists? Despite all the fuss about the science of foreign policy, the US experience shows clearly that another type of expertise has proved crucial. Faculties have developed huge area study programmes. The basic rationale was obvious. In order to better understand the world, people learnt the language and culture of other countries, coupled with some social science tools. Not some ready-made scientific laws, but some contextual knowledge was the answer.

Here, teaching the constitutive function of theories and, more precisely,

teaching how different understandings might exist at the same time can fulfil a crucial learning function, similar to that in learning languages/culture: it prepares one for cross-cultural understanding and communication. To some extent, learning *via* philosophy/meta-theory different theoretical languages can substitute some of the heuristic functions of learning real languages in cultural studies.¹³ Such an approach can have an economic appeal as well: in the absence of the means necessary to fund expensive international field studies and area specialisations, one simply needs some books — and mainly brains.

Moreover, teaching theory contributes to developing negotiating skills. For students are led to think as if they were explaining something from different points of view. They have to put themselves into a theoretical frame and produce explanations accordingly. Most importantly, when discussing with other students who might not share their interpretation, they have to learn how to make their own argument palatable to their opponent, to translate their ideas into the theoretical language of the other. It makes students able to decipher the other's position in terms of their assumption, and to respond by using this knowledge. They can therefore anticipate reactions more quickly.

Such skills make for a more reflexive process of communication that should reduce misunderstandings. These are hermeneutical skills that are crucial for the observing and acting diplomat. If acquired, they are more time-proven and will be useful for any practitioner when some other form of knowledge has become outdated. In that regard, theoretical discussions are fundamental for developing the intellectual maturity of future actors in international affairs.

The golden rule of diplomacy is not to impose one's visions on the opponent but

The
Significance
and Roles of
Teaching
Theory in
International
Relations

Stefano
Guzzini

VKubalkova
2004-01-04 12:55:30

read it up to her please

to change their preferences so as to make them compatible with one's own. This assumes good hermeneutical skills. Similarly, Kissinger used to complain about the recruitment of lawyers and business people to the State Department, exactly because they lacked this capacity to translate into different environments taking the world as one. Problems are treated *ad hoc* with little historical distance: 'Nations are treated as similar phenomena and those states presenting similar immediate problems are treated similarly' (Kissinger 1969:33). Instead, Kissinger asked for more historical but also

knowledge. He called for technical understanding requires an intellectual self achieved through his which was on the "mean- nothing less), that is, in not philosophy

Academic Getting Out phery

is obvious yet perhaps argument in favour theory in IR has to do arch agenda in IR. To put it crudely: IR (and increasingly more specialised studies on the European Union) has become a fully-fledged social science whose debates are basically driven by theory. Contributions are judged on their ability to advance more general knowledge that can be made accessible to a wider audience. Theory is part and parcel of all articles in the leading journals of the field: *International Organization*, *International Security*, *International Studies Quarterly*, and *World Politics* (for the US), as well as the *European Journal of International Relations*, the *Journal of Peace Research*, *Millennium*, the *Review of*

International Studies, and the new *Zeitschrift für Internationale Beziehungen*. And this seems justified in the light of the arguments given above.

Theoretical strength has become an indicator of defining an academic core and an academic periphery. Let Germany and Italy serve as examples. Germany's academic IR community, although quite numerous, had no single anonymously peer-reviewed journal in its field until 1994. By that time, the IR section of the German Political Science Association decided to get rid of some of the classical feudal features of any purely cooptative, rather than meritocratic, academic system. It created the *Zeitschrift für Internationale Beziehungen*, which has met with an impressively quick and resounding success. Big shots see their articles refused. Lesser known figures become part of a more general debate. Whereas the traditional IR elite was already quite open to theoretical research (e.g. Czempel, Senghaas, Krippendorff),¹⁴ by now the journal displays articles of a theoretical density fully competitive with any major international journal – a richness sustained by the mainly younger generations in the field. The German debate has moved out of the semi-periphery in which it found itself for quite a time. It did so by creating an internal intellectual dynamic and by connecting its findings to the core of research abroad. There was no way to do this without independent research on theory.¹⁵

This is not quite the situation of IR in Italy. In a recent study, Lucarelli and Menotti (2000) draw a rather critical picture of the Italian scene. There is, relatively speaking, little theoretical work in its widest sense, i.e. including those works which explicitly link theoretical models to empirical material. The most remarkable data is, perhaps, the number of Italian authors who have published in

leading international journals. Defining an Italian scholar as one who works in Italy or is part of its academic community, there are only two people who have published in total four articles. For comparison, it would be easy to fill pages with German or Scandinavian names. Although there are journals of good, if mixed, quality like *Teoria Politica*, there is non really anonymously and internationally peer-reviewed journal. The debates go through the big names.

Now one could say that it is not necessarily a sign of being peripheral if the academic community does not publish in international journals. Some academic communities that used to happily publish at home only (given that their internal circles were big enough) were very much part of the international stage, because other communities referred to their work. French philosophy and sociology could serve as examples. Even if we grant that this could be a temporary possibility it is, however, quite natural to resume the link to the international level — a link which has passed necessarily *via* theoretical exchanges. Having no link for quite some time, in either direction, like in Italy, still seems to be a problem.

Still, one could reply that it is perhaps not disastrous if some academic communities do not torture their brains by following the latest intricacies of the “relative *versus* absolute gains” debate or similar theoretical research in IR. An academic periphery might serve as a good protective shield, in the way that poverty was often the best monument preserver during the European Baroque. There is undoubtedly some truth to this. Yet after so many years one needs something worth preserving, something up to the standards of later times. Moreover, moving out of the periphery does not imply that one has to imitate everything. As Zürn (1994) so nicely put it in the opening issue

of the *Zeitschrift für Internationale Beziehungen*, ‘We can do much better! *Aber muß es auf amerikanisch sein?*’ Both the dynamic of the Scandinavian and German debates have rejuvenated a Continental debate which, although obviously neglected in the US (Waever 1998), is now well alive and kicking (Jørgensen 2000). This renewal has been accompanied and reinforced by the establishment of a new (since 1995) and internationally peer-reviewed journal of very high quality, the *European Journal of International Relations*, first edited in Uppsala (Carlsnaes) and now in München (Kratochwil).

All of these points seem perhaps trivial to many, but these experiences are important for those countries which are still defining their place in the international division of academic labour, such as the new, and sometimes not so new, academic communities in Central and Eastern Europe. And here the picture seems often similar if not worse than in Italy. Whereas theoretical panels are very prominent in the Western European conference life, such as in the pan-European (well ...) ECPR (European Consortium for Political Research) Standing Group on IR meetings, only the last conference of the CEEISA (Central and East European International Studies Association) included some theoretical research.

Some of the reasons mentioned for this remarkable lacuna are that young or poor countries must focus on their most immediate technical needs. It seems undeniable that some new technical expertise is warranted. I hope that the two earlier sections have shown that any neglect of theoretical work would be short-sighted, though. Moreover, since future teachers have to be trained, theoretical requirements are even more important since, without it, no academic work (including Ph.D.s!) of an international standard can be expected.

The
Significance
and Roles of
Teaching
Theory in
International
Relations

Stefano
Guzzini

Indeed, there is a risk in the relative comfort the semi-periphery provides. Academic communities, for instance in Central and Eastern Europe, might simply accept the position in which the international division of academic labour will try to slot them, namely one in which they teach and research only on their particular region, passively relying on theories invented by somebody else (or worse, without even any theoretical background which would make them able to relate to other phenomena). The risk is great that scholars will be content with filling out those chapters where "regional expertise" is needed. "The view from ..." litters book chapter headings like titles in United Nations reports. Ph.D. theses will be guided by these requirements mainly. Even in Germany, many Ph.D.s are financed through research projects that rarely extend to fundamental research. As a result, quite a few young Germans see little other way than to graduate in the United Kingdom or the US.

Such a position risks cementing the semi-periphery. Let us, for the sake of drama, express this in crudely economic terms. Knowledge follows, to some extent, a similar path as other products in international trade. Countries are, of course, free to specialise in raw materials, but the history of international trade has shown that there are limits to this. Exchanging their goods, these countries have come to know a dependence on international technology and tastes. Usually their prices are driven down compared to high value-added goods with a high knowledge component. In particular, Japan stands out as a country which found out that it is simply not enough to copy things, as one must also understand the logic of production. Instead of being a technology-taker, one should be a trend-setter in new technologies. Know-how also derives from basic science, hence the

latter simply cannot be disregarded, neither in telecommunications, nor in academic production. Of course, that takes time. But it does not happen if one never starts. Without acknowledging the need for theory, and without developing the possibility for theoretical studies to develop, academic communities risk staying or becoming simple theory-takers (*i.e.* passive knowledge consumers) and mere data-providers.

A Practical Note on Teaching IR Theory

This conception of the significance and role of theories in IR is not inconsequential for its actual teaching. There are some obvious points. Since all empirical analysis, and all history, implies theoretical assumptions, one really cannot do applied studies first and theory later. Consequently, it makes little sense to wait for meta-theory/theory-content in courses of later semesters, although the type of course may differ. Also, since seminars are to be geared towards perspectivist thinking and theory-translation, the seminar leaders themselves must be competent in a variety of different approaches. In a similar vein, instead of a curriculum in which there has been a division of theoretical spheres of influence with little exchange in-between, it is preferable to have several courses which in themselves try to make people think on the basis of a variety of approaches, although the preference of the seminar leaders will differ.

Since the foregoing discussion has stressed the importance of the constitutive function of theories, I will not deal with the majority of courses that have a necessary theoretical content, but which are not specifically geared towards preparing students for theoretical think-

ing.¹⁶ Instead, I will focus on those courses that more explicitly tackle the constitutive function of theories. These more theoretical courses are concerned with the way explanations are constructed. They must discuss the assumptions upon which middle-range explanatory theories are built. In other words, these courses must discuss meta-theories (of course not exclusively). There is no finite list for teaching the inter-relationship between meta-theories and theories in IR. Also, teaching should follow the needs of particular students and there is little generalisation that can be offered. Still, some patterns can be discerned.

There are perhaps four basic ways of combining theory/meta-theory in a course. The most classical way, at least on the European continent, would be *via* a history of thought. Strangely enough, this is rarely done. There are two basic strategies for such an approach. One can either refer to the philosophical forerunners of international theories, or limit oneself to the theoreticians that were prominent in the field that became institutionalised as a discipline after the World War I. In Western dominated IR, such a course then tends to be structured around the so-called four great debates (realism-idealism, scientism-traditionalism, realism-globalism, rationalism-constructivism), but that can be enriched by local references. The advantage of such an approach is that it introduces some of the peculiar IR language to students (security dilemma, balance of power, national interest), all its pitfalls, as well as some methodological issues in a "chronological" manner. Moreover, if presented in a sociological way, students get some insight into the relationship between world political events and theoretical advances, *i.e.* they learn to put the production of knowledge into its social and historical contexts. The disadvantage is that the core will inevitably centre around real-

ism in IR, and will hence not present the variety of approaches (as seen also in Guzzini 1998). Moreover, it can meet the incomprehension of students who might perceive this as an exercise in resurrecting mummified ideas, a kind of intellectual archaeology. Hence, it is important that links to contemporary affairs or debates be made throughout the course, something which is actually not that difficult.

A second approach, which works in a more straightforward analytical way, was prominent in the 1980s and 1990s, and not only in the West. It consists of presenting theory as a menu for choosing between clusters of assumptions bundled as schools of thought or paradigms. Let us call it the approach of the 'Inter-Paradigm Debate' (Banks 1985). There are some relatively famous books on realism, pluralism and globalism or similar titles.¹⁷ This also applied to International Political Economy where, again, we were usually offered three choices: mercantilism, liberalism and structuralism/neo-Marxism (Gilpin 1975; Gilpin, with Gilpin 1987; Gill 1988). Such a didactic approach has some obvious advantages. It is logical in that it takes first what comes first: the underlying assumptions of all theories and observation. Also, it immediately does some homework for the student in that it shows that assumptions usually come in a cluster. Finally, it trains students explicitly in method. The usual game consists of asking students to explain an event in any of the theories, which means that students must understand how hypotheses are formulated and put up for empirical scrutiny.

There are a series of caveats to this second approach, however. One is the above-mentioned and often encountered confusion of ideologies with meta-theories. In other words, there exists the risk of taking these clusters in purely ideological terms. Yet, one should be well advised

The
Significance
and Roles of
Teaching
Theory in
International
Relations

Stefano
Guzzini

not to confound realism/pluralism/structuralism with conservatism/liberalism/radicalism, as often done especially in the US (Gilpin, with Gilpin 1987). If conceived in such an ideological way, the course might have the opposite effect: instead of opening up for thinking, it closes down the path to debate. If diverging values are all there is, then the debate can easily turn into a show of verbal fists. Worse, in some settings, the intellectual exchange might never start since everybody feels entitled to stick to what they think anyway (and professors are always right). At this point, theory courses simply add one more to the list of compulsory hurdles later quickly forgotten. Also, students do not necessarily get the link between knowledge and the social environment in which it is produced (the mirror argument to the advantage of the first more historical technique). So the "Inter-Paradigm Debate" was perhaps best understood as a historical phenomenon, *i.e.* as an indicator of the self-understanding of a discipline in crisis, and less as a logical clash of incommensurable paradigms (Guzzini 1988; 1998).

The present discussion between rationalism and constructivism is preferable to the classical Inter-Paradigm Debate in that it focuses very well on the meta-theoretical differences.¹⁸ In other words, it does not reduce meta-theory to an ideological factor. The meta-theoretical matrix on which it is built has, however, inconveniences of its own. The two most important are the fuzzy borders of some categories and the general evacuation of all normative debates (Guzzini 1998:190-235). Moreover, the disadvantage of teaching in this way is the "top-down" manner. Students previously unacquainted with the need to get some distance from the data presented to them will find it hard to make the theoretical-empirical link. Of course, teaching can help by using articles

as examples, and so on. Similarly, small assignments that repeatedly ask to generate and compare explanations on the basis of different theories are helpful. Finally, newspaper articles can be used and compared to tease out the underlying assumptions. All this is relatively time-intensive, though, since the instructor has to give very thorough and frequent feedback to students. Hence, depending on the size of the class, the tasks of communication and learning control can prove difficult. To put it differently, such theory courses can hardly be lectures and do not work well in big settings.

A third way of teaching the constitutive function of theories is to take a central concept, like "power", "security", "world society" or "war", and then show how the meaning of such concepts, as well as their explanatory value, diverges from one meta-theoretical/theoretical context to another. For such central concepts, once they are made part of the vocabulary of an observer (applied or theoretical) they derive their significant meaning from the contexts in which they are embedded (Guzzini 1993). Also, depending on the concepts, this is a good way to refer to normative debates in IR.

The limits of such courses are similar to those met by the fourth type of enhancing the awareness of the constitutive function of theories, which explicitly starts from the empirical end. One way, for instance, would be to take an important international event, as *e.g.* the above-mentioned "end of the Cold War". The seminar confronts different interpretations of the event. Then, it discusses the relative value of these interpretations on different levels: their empirical accuracy, their internal coherence, and finally the plausibility of the assumptions, both open and hidden, on which the explanation is based. The aim is to get students

interested in exploring these assumptions, and hence to ask theoretical questions. This way works by motivating students to think theoretically starting from an empirical theme.

The drawbacks of this way of teaching are that assumptions are hardly seen in context or in a bundle. It leaves a certain taste of *ad-hoc*-ness to it. Again, the teacher can try to remedy this by referring to the larger theories to which particular assumptions tend to belong. Without proper preparation, however, this strategy risks moving the discussion too early to a higher theoretical level. Still, such a course could be good at the early stages of a student life, trying to stimulate the necessary theoretical awareness – and might then be repeated in a more advanced way at a later stage.

Conclusion

This article has sought to oppose the often-encountered conflation between theory and academia on the one hand, and politics and practice, on the other. Indeed, it has tried to show that these distinctions are themselves particularly damaging since they reinforce the illusion that empirical studies can ever be divorced from theory, and that education in international affairs is a purely “professional” exercise that stresses mere knowledge and not reflexive skills. In other words, these distinctions hide the specific contributions of theory training for practitioners in IR.

I advanced three arguments for better assessing the significance and roles of teaching theory in IR. First, all empirical analysis implies theory. Theory is not something to be added later. Its awareness is not an “academic” enterprise, but actual politics. Second, in times of changing knowledge, the teaching of theory

can help foster a more time-independent self-reflexivity and cultural and contextual awareness, in short, a form of intellectual maturity particularly useful for cross-national understanding, communication, and negotiation. Finally, and perhaps also for this reason, teaching and researching in theory, including fundamental research, has been a major ingredient in moving academic communities out of the (semi-)periphery.

I ended with a short note on actual teaching by identifying four classical types of courses. This was not meant to exhaust all the possible ways to raise the hermeneutical skills deemed necessary for future practitioners (in academia, journalism, economics or politics). Also, they should not suggest that this kind of course is all there should be. To the contrary, as mentioned above, it cannot replace a good training in basic processes and issues in international history, law, politics and political economy. Yet, as I simply wanted to stress, neither can applied studies replace such theoretical courses, relying perhaps more on sociology and philosophy. Nor does it appear that they would be “more important” to the future practitioners who will, as academics should have no trouble admitting, necessarily learn a good part of their applied trade while doing it.

There is perhaps an even more far-reaching implication of the argument. The call for more applied studies betrays a longing for a time when academia and politics spoke basically the same language. In the past, realism, or more generally the classical tradition, has been the bridge between observation and decision. This bridge was taken for granted, the language was natural. There was an easy coming and going between scholars and politicians. Today, the impression is that the language of academia has increasingly moved away from that of practice. Hence,

The
Significance
and Roles of
Teaching
Theory in
International
Relations

the call for re-importing the language of politics into academia.

But such a resurrection of the old united language is impossible to achieve in some countries by now. As practitioners have noticed themselves, there is no way back to the "natural" language of scholars and practitioners. It is this very self-awareness of needing competent translations that makes the return to a *status quo* very difficult. The realist or classical language and view of international affairs is no longer obvious in many countries. That is because realism or the parameters of the realist-idealist debate have in themselves become an object of study, exposing it as a set of practices which in itself influences the reality it is supposed to passively explain.

Therefore, it is also contestable whether this rapprochement should be done in this conservative way. Resisting this rising self-awareness and the exposure to academic distance with the excuse that the languages no longer fit is not an innocent move. It would have as an effect that the future elite, which might have been trained in these new professionalised academic environments, would itself stop speaking the old language of politics. For all the above does not imply that the link between academia and practice is lost forever. Rather, it has to be redefined. Practitioners can be self-reflective persons who are able to combine the old and new languages of practice and of observation, as diplomats have always been speaking different languages to different audiences. Academics should be able to include in the analysis the self-fulfilling effects that certain explanations can have. True, this implies a double socialisation. But instead of pursuing the conservative endeavour to paper over a unity lost, one might face the challenge of redefining it. It would not be difficult for constructivists to live with this commit-

ment to self-reflexivity (Guzzini 2000a). The significance and roles of theory in teaching IR and researching international relations would be no matter for dispute.

First version received: July 2000.

Final version accepted: March 2001.

Notes:

Stefano Guzzini is Senior Research Fellow at the Copenhagen Peace Research Institute (COPRI), and Associate Professor of Political Science, International Relations and European Studies at the Central European University, Budapest (on leave).

Address: Stefano Guzzini, Copenhagen Peace Research Institute, Fredericiagade 18, DK-1310 Copenhagen K, Denmark [E-mail: sguzzini@copri.dk].

This is a revised version of a paper presented at the training seminar 'Teaching European Studies and International Relations in East Central Europe' organised by the UNESCO Chair in International Relations and European Studies at the Institute of International Relations and Political Science, Vilnius University, 2-5 March 2000. For comments and suggestions, the author is indebted to the participants of this workshop, in particular Raimundas Lopata, Nortautas Statkus, and Gedeminas Vitkus, as well as to Kateřina Borutová, Barry Buzan, Lene Hansen, Ulla Holm, Pertti Joenniemi, Işıl Kazan, Anna Leander, Ian Manners, Michael Merlingen, Karen Lund Petersen, Ulrich Sedelmeier, Ole Waever, Zlatko Čabić and two anonymous referees.

¹ For recent assessments of the findings, see Vasquez (1987), Geller and Singer (1998).

² Indeed, one should not forget that empiricism is a highly sceptical position which was born out of the criticism to derive knowledge from preconceived ideologies, instead of what there "really is" which, in turn, is what our senses, and only those, tell us there is. In the context of heightened ideological debate, like the Cold War, it was only to be expected that peace researchers found empiricist research methods attractive. Unfortunately, no amount of data

would have been enough to undermine the faith in *Realpolitik* (Vasquez 1983; 1998).

3 For a longer discussion, see Guzzini (2000a).

4 There are libraries written on this by now. Yet on the theoretical debate, see Lebow and Stein (1994), Wohlforth (1994-95; 1998), Lebow and Risse-Kappen (1995), Forsberg (1999), Lebow (1999), Patman (1999), Schweller and Wohlforth (2000), Brooks and Wohlforth (2000-01), as well as the exchange between Kramer (1999; 2001), and Wohlforth (2000).

5 In reality there is little debate about it, at least not in IR. Perhaps this simply shows the lack of trained historians in the field.

6 The more profound argument is that "power" is an indeterminate concept which cannot play the same role as money in economic theory. For that argument, see already Aron (1962:98), and Wolfers (1962:196). See also Baldwin (1989:25, 209) whose conceptual analysis shows that overall concepts of power, as used in classical balance of power theories are 'virtually meaningless'. For a discussion, see Guzzini (1993; 2000b).

7 For a US version, see Vasquez (1983).

8 See the exchange collected in Lynn-Jones (1991).

9 Much earlier, geo-politicians heralded population exchange as a rational tool for conflict resolution. See, for instance, Mackinder (1919/1944) reflecting on the Turkish-Greek treaty in 1923.

10 See the debate between Bueno de Mesquita (1985) and Krasner (1985).

11 See the long development of a conceptual framework which precedes his *Economy and Society* (Weber 1922/1980).

12 For this argument, see also Wendt's (1999:83-8) discussion about the difference between causal and constitutive theories and the importance of conceptual analysis.

13 Theoretical training can of course not entirely

replace the exposure to different world-views and cultures. As such, it is relatively sad that many top politicians of the leading power of the day are not exactly known for their expertise in foreign languages and cultures. But theoretical training goes hand in hand with that exposure: being trained in the constitutive function of theories prepares the ground for a better understanding and use of the exposure to different cultures, exactly because there is a greater awareness of one's own values.

14 As an example for a much wider body of literature, see Senghaas (1972; 1987), Krippendorff (1975; 1977; 1985) and Czempel (1981; 1989).

15 These theoretical contributions can be found (1) in the original debate between rational choice and communicative action in IR (a different take on the debate on rationalism and constructivism) by Meyers (1994), Müller (1994; 1995), Schneider (1994), Keck (1995; 1997), Risse-Kappen (1995), Schmalz-Bruns (1995) and now exported to the US by Risse (2000); see also the contributions to constructivism by Jaeger (1996) and Zehfuss (1998) and on action theory by Schimmelfennig (1997); (2) in the debate around democratic peace theory by Risse-Kappen (1994), Czempel (1996a; 1996b), Moravcsik (1996) and Schmidt (1996); (3) or in the debate about globalisation and post-national politics (e.g. Brock and Albert 1995; Zürn 1997; Jung 1998; Schmalz-Bruns 1999; and the special issue by Grande and Risse 2000).

16 For instance, in a course on conflict resolution, one can teach middle-range theories that specify the conditions under which attempts at third-party mediation might be more successful.

17 For surveys of this triad, see Rittberger and Hummel (1990:23), and Waever (1996:153).

18 For the classical exposition, see Keohane (1988/1989) and Lapid (1989); for its present status, see Katzenstein *et al.* (1998) and Guzzini (2000a).

The
Significance
and Roles of
Teaching
Theory in
International
Relations

References

- Aron, Raymond (1962) *Paix et la guerre entre les nations*. Paris: Calmann-Lévy.
Baldwin, David A. (1989) *Paradoxes of Power*. Oxford: Blackwell.
Banks, Michael (1985) The Inter-Paradigm Debate. In Margot Light and A. J. R. Groom (eds) *International Relations: A Handbook of Current Theory*, 7-26. London: Frances Pinter.

- Brock, Lothar and Mathias Albert (1995) Entgrenzung der Staatenwelt. Zur Analyse weltgesellschaftlicher Entwicklungstendenzen. *Zeitschrift für Internationale Beziehungen* 2(2), 259-85.
- Brooks, Stephen G. and William C. Wohlforth (2000-01) Power, Globalization, and the End of the Cold War: Reevaluating a Landmark Case for Ideas. *International Security* 25(3), 5-53.
- Brown, Chris (1997) *Understanding International Relations*. Houndmills and London: Macmillan.
- Bueno de Mesquita, Bruce (1985) Toward a Scientific Understanding of International Conflict: A Personal View. *International Studies Quarterly* 29(1), 121-36.
- Collier, David and James Mahon (1993) Conceptual 'Stretching' Revisited: Adapting Categories in Comparative Politics. *American Political Science Review* 87(4), 845-55.
- Czempiel, Ernst-Otto (1981) *Internationale Politik. Ein Konfliktmodell*. Paderborn: Schöningh.
- Czempiel, Ernst-Otto (1989) Internationalizing Politics: Some Answers to the Question of Who Does What to Whom. In Ernst-Otto Czempiel and James Rosenau (eds) *Global Changes and Theoretical Challenges: Approaches to World Politics for the 1990s*, 117-34. Lexington, MA: Lexington Books.
- Czempiel, Ernst-Otto (1996a) Kants Theorem. Oder: Warum sind die Demokratien (noch immer) nicht friedlich? *Zeitschrift für Internationale Beziehungen* 3(1), 79-101.
- Czempiel, Ernst-Otto (1996b) Theorie und Strategie. Überlegungen nach Hajo Schmidts Kommentar. *Zeitschrift für Internationale Beziehungen* 3(1), 117-122.
- Fontaine, André (1965) *Histoire de la Guerre Froide. Tome I: De la révolution d'octobre à la guerre de Corée, 1917-1950*. Paris: Fayard.
- Forsberg, Tuomas (1999) Power, Interests and Trust: Explaining Gorbachev's Choices at the End of the Cold War. *Review of International Studies* 25(4), 603-21.
- Geller, Daniel S. and David J. Singer (1998) *Nations at War: A Scientific Study of International Conflict*. Cambridge: Cambridge University Press.
- Gill, Stephen (1988) *The Global Political Economy*. Brighton: Harvester Wheatsheaf.
- Gilpin, Robert (1975) Three Models of the Future. *International Organization* 29(1), 37-60.
- Gilpin, Robert, with the assistance of Jean M. Gilpin (1987) *The Political Economy of International Relations*. Princeton, NJ: Princeton University Press.
- Gilpin, Robert, with the assistance of Jean M. Gilpin (2000) *The Challenge of Global Capitalism: The World Economy in the 21st Century*. Princeton, NJ: Princeton University Press.
- Grande, Edgar and Thomas Risse, eds (2000) Globalisierung und die Handlungsfähigkeit des Nationalstaates. Special issue of *Zeitschrift für Internationale Beziehungen* 7(2).
- Grosser, Alfred (1972) *L'explication politique: une introduction à l'analyse comparative*. Paris: Armand Colin.
- Guzzini, Stefano (1988) T. S. Kuhn and International Relations: International Political Economy and the Inter-Paradigm Debate. London: London School of Economics and Political Science, MSc (Econ) Thesis.
- Guzzini, Stefano (1993) Structural Power: The Limits of Neorealist Power Analysis. *International Organization* 47(3), 443-78.
- Guzzini, Stefano (1998) *Realism in International Relations and International Political Economy: The Continuing Story of a Death Foretold*. London and New York: Routledge.
- Guzzini, Stefano (2000a) A Reconstruction of Constructivism in International Relations. *European Journal of International Relations* 6(2), 147-82.
- Guzzini, Stefano (2000b) The Use and Misuse of Power Analysis in International Theory. In Ronen Palan (ed.) *Global Political Economy: Contemporary Theories*, 53-66. London and New York: Routledge.
- Huntington, Samuel P. (1993) The Clash of Civilizations? *Foreign Affairs* 72(3), 22-42.
- Jaeger, Hans-Martin (1996) Konstruktionsfehler des Konstruktivismus in den Internationalen Beziehungen. *Zeitschrift für Internationale Beziehungen* 3(2), 313-40.
- Jørgensen, Knud-Erik (2000) Continental IR Theory: The Best Kept Secret. *European Journal of International Relations* 6(1), 9-42.

Stefano
Guzzini

- Jung, Dietrich (1998) Weltgesellschaft als theoretisches Konzept der Internationalen Beziehungen. *Zeitschrift für Internationale Beziehungen* 5(2), 241-71.
- Katzenstein, Peter, Robert O. Keohane and Stephen D. Krasner (1998) International Organization and the Study of World Politics. *International Organization* 52(4), 645-85.
- Keck, Otto (1995) Rationales kommunikatives Handeln in den internationalen Beziehungen. Ist eine Verbindung von Rational-Choice-Theorie und Habermas' Theorie des kommunikativen Handelns möglich? *Zeitschrift für Internationale Beziehungen* 2(1), 5-48.
- Keck, Otto (1997) Zur sozialen Konstruktion des Rational-Choice Ansatzes. Einige Klarstellungen zur Rationalismus-Konstruktivismus Debatte. *Zeitschrift für Internationale Beziehungen* 4(1), 139-51.
- Keohane, Robert O. (1988/1989) International Institutions: Two Approaches. In Robert O. Keohane (ed.) *International Institutions and State Power: Essays in International Relations Theory*, 158-79. Boulder, CO: Westview Press.
- Keohane, Robert O. (1989) *International Institutions and State Power: Essays in International Relations Theory*. Boulder, CO: Westview Press.
- Kindleberger, Charles P. (1973/1987) *The World in Depression, 1929-1939*. Berkeley, CA: University of California Press.
- King, Gary, Robert O. Keohane and Sidney Verba (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.
- Kissinger, Henry A. (1969) *American Foreign Policy: Three Essays*. New York: Norton & Co.
- Koslowski, Roy and Friedrich Kratochwil (1994) Understanding Change in International Politics: The Soviet Empire's Demise and the International System. *International Organization* 48(2), 215-47.
- Kramer, Mark (1999) Ideology and the Cold War. *Review of International Studies* 25(4), 539-76.
- Kramer, Mark (2001) Realism, Ideology and the End of the Cold War: A Reply to William Wohlforth. *Review of International Studies* 27(1), 119-30.
- Krasner, Stephen D. (1985) Toward Understanding in International Relations. *International Studies Quarterly* 29(1), 137-44.
- Krippendorff, Ekkehart (1975) *Internationales System als Geschichte*. Frankfurt/M.: Campus.
- Krippendorff, Ekkehart (1977) *Internationale Beziehungen als Wissenschaft*. Frankfurt/M.: Campus.
- Krippendorff, Ekkehart (1985) *Staat und Krieg. Die historische Logik politischer Unvernunft*. Frankfurt/M.: Suhrkamp.
- Lapid, Yosef (1989) The Third Debate: On the Prospects of International Theory in a Post-Positivist Era. *International Studies Quarterly* 33(3), 235-54.
- Lebow, Richard Ned (1999) The Rise and Fall of the Cold War. *Review of International Studies* 25(special issue), 21-39.
- Lebow, Richard Ned and Janice Gross Stein (1994) *We All Lost the Cold War*. Princeton, NJ: Princeton University Press.
- Lebow, Richard Ned and Thomas Risse-Kappen, eds (1995) *International Relations Theory and the End of the Cold War*. New York: Columbia University Press.
- Lucarelli, Sonia and Roberto Menotti (2000) 'IR Theory in Italy in the 1990s'. Copenhagen: ECPR, 28th Joint Sessions of Workshops, 14-19 April.
- Lynn-Jones, Sean M., ed. (1991) *The Cold War and After: Prospects for Peace*. Cambridge, MA: The MIT Press.
- Mackinder, Halford John (1919/1944) *Democratic Ideals and Reality: A Study in the Politics of Reconstitution*. Harmondsworth: Penguin.
- McGwire, Michael (1991) *Perestroika and Soviet National Security*. Washington, DC: The Brookings Institution.
- Mearsheimer, John (1990) Back to the Future: Instability in Europe After the Cold War. *International Security* 15(1), 5-56.
- Meyers, Reinhard (1994) Virtuelle Scheingefechte im ontologischen Cyberspace? Nachfragen zum Duktus und zum Gehalt einer Theoriedebatte. *Zeitschrift für Internationale Beziehungen* 1(1), 127-37.

The
Significance
and Roles of
Teaching
Theory in
International
Relations

- Moravcsik, Andrew (1996) Federalism and Peace: A Structural Liberal Perspective. *Zeitschrift für Internationale Beziehungen* 3(1), 123-32.
- Morgenthau, Hans J. (1948) *Politics Among Nations: The Struggle for Power and Peace*. New York: Knopf.
- Müller, Harald (1994) Internationale Beziehungen als kommunikatives Handeln. Zur Kritik der rationalistischen Handlungstheorien. *Zeitschrift für Internationale Beziehungen* 1(1), 15-44.
- Müller, Harald (1995) Spielen hilft nicht immer. Die Grenzen des Rational-Choice-Ansatzes und der Platz der Theorie kommunikativen Handelns in der Analyse internationaler Beziehungen. *Zeitschrift für Internationale Beziehungen* 2(2), 371-91.
- Patman, Robert G. (1999) Reagan, Gorbachev and the Emergence of 'New Political Thinking'. *Review of International Studies* 25(4), 577-601.
- Richter, Rudolf (1994) Methodology From the Viewpoint of the Economic Theorist – Thirty Years On. *Journal of International and Theoretical Economics* 150(4), 589-608.
- Stefano Guzzini Risse, Thomas (2000) 'Let's Argue!': Communicative Action in World Politics. *International Organization* 54(1), 1-39.
- Risse-Kappen, Thomas (1994) Wie weiter mit dem "demokratischen Frieden"? *Zeitschrift für Internationale Beziehungen* 1(2), 367-79.
- Risse-Kappen, Thomas (1995) Reden ist nicht billig. Zur Debatte um Kommunikation und Rationalität. *Zeitschrift für Internationale Beziehungen* 2(1), 171-84.
- Rittberger, Volker and Hartwig Hummel (1990) Die Disziplin 'Internationale Beziehungen' im deutschsprachigen Raum auf der Suche nach ihrer Identität: Entwicklung und Perspektiven. In Volker Rittberger (ed.) *Theorien der Internationalen Beziehungen. Bestandsaufnahme und Forschungsperspektiven*, 17-47. Opladen: Westdeutscher Verlag.
- Sartori, Giovanni (1970) Concept Misformation in Comparative Politics. *American Political Science Review* 64(4), 1033-53.
- Schimmelfennig, Frank (1997) Rhetorisches Handeln in der internationalen Politik. *Zeitschrift für Internationale Beziehungen* 4(2), 219-54.
- Schmalz-Bruns, Rainer (1995) Die Theorie kommunikativen Handelns – eine Flaschenpost? *Zeitschrift für Internationale Beziehungen* 2(2), 347-370.
- Schmalz-Bruns, Rainer (1999) Deliberativer Supranationalismus. Demokratisches Regieren jenseits des Nationalstaats. *Zeitschrift für Internationale Beziehungen* 6(2), 185-244.
- Schmidt, Hajo (1996) Kant und die Theorie der Internationalen Beziehungen. Vom Nutzen und den Problemen einer aktualisierten Kantlektüre – ein Kommentar zu E.-O. Czempiel. *Zeitschrift für Internationale Beziehungen* 3(1), 103-16.
- Schneider, Gerald (1994) Rational Choice und kommunikatives Handeln. Eine Replik auf Harald Müller. *Zeitschrift für Internationale Beziehungen* 1(2), 357-66.
- Schweller, Randall L. and William C. Wohlforth (2000) Power Test: Evaluating Realism in Response to the End of the Cold War. *Security Studies* 9(3), 60-107.
- Senghaas, Dieter (1972) *Rüstung und Militarismus*. Frankfurt/M.: Suhrkamp.
- Senghaas, Dieter (1987) *Konfliktformationen im internationalen System*. Frankfurt/M.: Suhrkamp.
- Smith, Steve (1995) The Self-Images of a Discipline. In Ken Booth and Steve Smith (eds) *International Relations Theory Today*, 1-37. Oxford: Polity Press.
- Strange, Susan (1986) *Casino Capitalism*. Oxford: Basil Blackwell.
- Strange, Susan (1998) *Mad Money*. Manchester: Manchester University Press.
- Vasquez, John (1983) *The Power of Power Politics: A Critique*. London: Frances Pinter.
- Vasquez, John (1987) The Steps to War: Toward a Scientific Explanation of Correlates of War Findings. *World Politics* 40(1), 108-45.
- Vasquez, John (1998) *The Power of Power Politics: From Classical Realism to Neotraditionalism*. Cambridge: Cambridge University Press.

- Wæver, Ole (1996) The Rise and Fall of the Inter-Paradigm Debate. In Steve Smith, Ken Booth and Marysia Zalewski (eds) *International Theory: Positivism and Beyond*, 149-84. Cambridge: Cambridge University Press.
- Wæver, Ole (1998) The Sociology of a Not so International Discipline: American and European Developments in International Relations. *International Organization* **52**(4), 687-728.
- Wallace, William (1996) Truth and Power; Monks and Technocrats: Theory and Practice in International Relations. *Review of International Studies* **22**(4), 301-21.
- Weber, Max (1922/1980) *Wirtschaft und Gesellschaft. Grundriss der Verstehenden Soziologie*. Tübingen: J. C. B. Mohr (Paul Siebeck).
- Wendt, Alexander (1992) Anarchy is What States Make of It: The Social Construction of Power Politics. *International Organization* **46**(2), 391-425.
- Wendt, Alexander (1999) *Social Theory of International Politics*. Cambridge: Cambridge University Press.
- Wohlforth, William C. (1994-95) Realism and the End of the Cold War. *International Security* **19**(3), 91-129.
- Wohlforth, William C. (1998) Reality Check: Revising Theories of World Politics in Response to the End of the Cold War. *World Politics* **50**(4), 650-80.
- Wohlforth, William C. (2000) Ideology and the Cold War. *Review of International Studies* **26**(2), 327-31.
- Wolfers, Arnold (1962) *Discord and Collaboration: Essays on International Politics*. Baltimore, MD: The Johns Hopkins University Press.
- Zehfuss, Maja (1998) Sprachlosigkeit schränkt ein. Zur Bedeutung von Sprache in konstruktivistischen Theorien. *Zeitschrift für Internationale Beziehungen* **5**(1), 109-37.
- Zürn, Michael (1994) We Can Do Much Better! Aber muß es gleich auf amerikanisch sein? Zum Vergleich der Disziplin "Internationale Beziehungen" in den USA und in Deutschland. *Zeitschrift für Internationale Beziehungen* **1**(1), 91-114.
- Zürn, Michael (1997) "Positives Regieren" jenseits des Nationalstaates. Zur Implementation internationaler Umweltregime. *Zeitschrift für Internationale Beziehungen* **4**(1), 41-68.

The
Significance
and Roles of
Teaching
Theory in
International
Relations

SOCIAL SCIENCE RESEARCH: PRINCIPLES, METHODS, AND PRACTICES

Anol Bhattacharjee, Ph.D.

University of South Florida
Tampa, Florida, USA
abhattach@usf.edu

Second Edition
Copyright © 2012 by Anol Bhattacharjee

Published under the Creative Commons Attribution 3.0 License



Chapter 4

Theories in Scientific Research

As we know from previous chapters, science is knowledge represented as a collection of "theories" derived using the scientific method. In this chapter, we will examine what is a theory, why do we need theories in research, what are the building blocks of a theory, how to evaluate theories, how can we apply theories in research, and also presents illustrative examples of five theories frequently used in social science research.

Theories

Theories are explanations of a natural or social behavior, event, or phenomenon. More formally, a scientific theory is a system of constructs (concepts) and propositions (relationships between those constructs) that collectively presents a logical, systematic, and coherent explanation of a phenomenon of interest within some assumptions and boundary conditions (Bacharach 1989).¹

Theories should explain why things happen, rather than just describe or predict. Note that it is possible to predict events or behaviors using a set of predictors, without necessarily explaining why such events are taking place. For instance, market analysts predict fluctuations in the stock market based on market announcements, earnings reports of major companies, and new data from the Federal Reserve and other agencies, based on previously observed *correlations*. Prediction requires only correlations. In contrast, explanations require *causations*, or understanding of cause-effect relationships. Establishing causation requires three conditions: (1) correlations between two constructs, (2) temporal precedence (the cause must precede the effect in time), and (3) rejection of alternative hypotheses (through testing). Scientific theories are different from theological, philosophical, or other explanations in that scientific theories can be empirically tested using scientific methods.

Explanations can be idiographic or nomothetic. **Idiographic explanations** are those that explain a single situation or event in idiosyncratic detail. For example, you did poorly on an exam because: (1) you forgot that you had an exam on that day, (2) you arrived late to the exam due to a traffic jam, (3) you panicked midway through the exam, (4) you had to work late the previous evening and could not study for the exam, or even (5) your dog ate your text book. The explanations may be detailed, accurate, and valid, but they may not apply to other similar situations, even involving the same person, and are hence not generalizable. In contrast,

¹ Bacharach, S. 8. (1989). "Organizational Theories: Some Criteria for Evaluation," *Academy of Management Review* (14:4), 496-515.

nomothetic explanations seek to explain a class of situations or events rather than a specific situation or event. For example, students who do poorly in exams do so because they did not spend adequate time preparing for exams or that they suffer from nervousness, attention-deficit, or some other medical disorder. Because nomothetic explanations are designed to be generalizable across situations, events, or people, they tend to be less precise, less complete, and less detailed. However, they explain economically, using only a few explanatory variables. Because theories are also intended to serve as generalized explanations for patterns of events, behaviors, or phenomena, theoretical explanations are generally nomothetic in nature.

While understanding theories, it is also important to understand what theory is not. Theory is not data, facts, typologies, taxonomies, or empirical findings. A collection of facts is not a theory, just as a pile of stones is not a house. Likewise, a collection of constructs (e.g., a typology of constructs) is not a theory, because theories must go well beyond constructs to include propositions, explanations, and boundary conditions. Data, facts, and findings operate at the empirical or observational level, while theories operate at a conceptual level and are based on logic rather than observations.

There are many benefits to using theories in research. First, theories provide the underlying logic of the occurrence of natural or social phenomenon by explaining what are the key drivers and key outcomes of the target phenomenon and why, and what underlying processes are responsible driving that phenomenon. Second, they aid in sense-making by helping us synthesize prior empirical findings within a theoretical framework and reconcile contradictory findings by discovering contingent factors influencing the relationship between two constructs in different studies. Third, theories provide guidance for future research by helping identify constructs and relationships that are worthy of further research. Fourth, theories can contribute to cumulative knowledge building by bridging gaps between other theories and by causing existing theories to be reevaluated in a new light.

However, theories can also have their own share of limitations. As simplified explanations of reality, theories may not always provide adequate explanations of the phenomenon of interest based on a limited set of constructs and relationships. Theories are designed to be simple and parsimonious explanations, while reality may be significantly more complex. Furthermore, theories may impose blinders or limit researchers' "range of vision," causing them to miss out on important concepts that are not defined by the theory.

Building Blocks of a Theory

David Whetten (1989) suggests that there are four building blocks of a theory: constructs, propositions, logic, and boundary conditions/assumptions. Constructs capture the "what" of theories (i.e., what concepts are important for explaining a phenomenon), propositions capture the "how" (i.e., how are these concepts related to each other), logic represents the "why" (i.e., why are these concepts related), and boundary conditions/assumptions examines the "who, when, and where" (i.e., under what circumstances will these concepts and relationships work). Though constructs and propositions were previously discussed in Chapter 2, we describe them again here for the sake of completeness.

Constructs are abstract concepts specified at a high level of abstraction that are chosen specifically to explain the phenomenon of interest. Recall from Chapter 2 that constructs may be unidimensional (i.e., embody a single concept), such as weight or age, or multi-dimensional (i.e., embody multiple underlying concepts), such as personality or culture. While some

constructs, such as age, education, and firm size, are easy to understand, others, such as creativity, prejudice, and organizational agility, may be more complex and abstruse, and still others such as trust, attitude, and learning, may represent temporal tendencies rather than steady states. Nevertheless, all constructs must have clear and unambiguous operational definition that should specify exactly how the construct will be measured and at what level of analysis (individual, group, organizational, etc.). Measurable representations of abstract constructs are called **variables**. For instance, intelligence quotient (IQ score) is a variable that is purported to measure an abstract construct called intelligence. As noted earlier, scientific research proceeds along two planes: a theoretical plane and an empirical plane. Constructs are conceptualized at the theoretical plane, while variables are operationalized and measured at the empirical (observational) plane. Furthermore, variables may be independent, dependent, mediating, or moderating, as discussed in Chapter 2. The distinction between constructs (conceptualized at the theoretical level) and variables (measured at the empirical level) is shown in Figure 4.1.

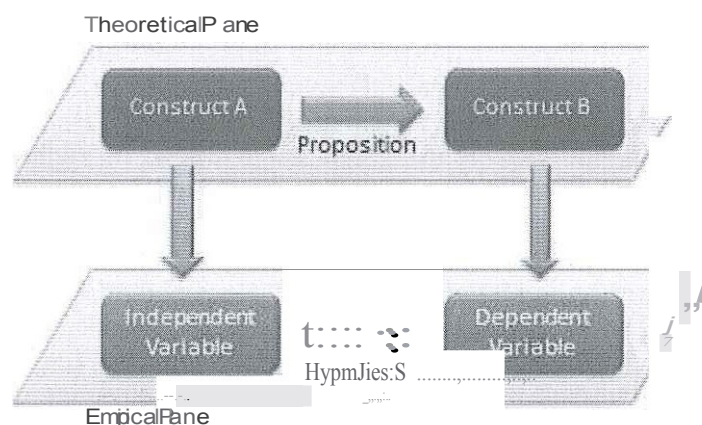


Figure 4.1. Distinction between theoretical and empirical concepts

Propositions are associations postulated between constructs based on deductive logic. Propositions are stated in declarative form and should ideally indicate a cause-effect relationship (e.g., if X occurs, then Y will follow). Note that propositions may be conjectural but **MUST** be testable, and should be rejected if they are not supported by empirical observations. However, like constructs, propositions are stated at the theoretical level, and they can only be tested by examining the corresponding relationship between measurable variables of those constructs. The empirical formulation of propositions, stated as relationships between variables, is called **hypotheses**. The distinction between propositions (formulated at the theoretical level) and hypotheses (tested at the empirical level) is depicted in Figure 4.1.

The third building block of a theory is the **logic** that provides the basis for justifying the propositions as postulated. Logic acts like a "glue" that connects the theoretical constructs and provides meaning and relevance to the relationships between these constructs. Logic also represents the "explanation" that lies at the core of a theory. Without logic, propositions will be ad hoc, arbitrary, and meaningless, and cannot be tied into a cohesive "system of propositions" that is the heart of any theory.

Finally, all theories are constrained by **assumptions** about values, time, and space, and **boundary conditions** that govern where the theory can be applied and where it cannot be applied. For example, many economic theories assume that human beings are rational (or

boundedly rational) and employ utility maximization based on cost and benefit expectations as a way of understanding human behavior. In contrast, political science theories assume that people are more political than rational, and try to position themselves in their professional or personal environment in a way that maximizes their power and control over others. Given the nature of their underlying assumptions, economic and political theories are not directly comparable, and researchers should not use economic theories if their objective is to understand the power structure or its evolution in an organization. Likewise, theories may have implicit cultural assumptions (e.g., whether they apply to individualistic or collective cultures), temporal assumptions (e.g., whether they apply to early stages or later stages of human behavior), and spatial assumptions (e.g., whether they apply to certain localities but not to others). If a theory is to be properly used or tested, all of its implicit assumptions that form the boundaries of that theory must be properly understood. Unfortunately, theorists rarely state their implicit assumptions clearly, which leads to frequent misapplications of theories to problem situations in research.

Attributes of a Good Theory

Theories are simplified and often partial explanations of complex social reality. As such, there can be good explanations or poor explanations, and consequently, there can be good theories or poor theories. How can we evaluate the "goodness" of a given theory? Different criteria have been proposed by different researchers, the more important of which are listed below:

- **Logical consistency:** Are the theoretical constructs, propositions, boundary conditions, and assumptions logically consistent with each other? If some of these "building blocks" of a theory are inconsistent with each other (e.g., a theory assumes rationality, but some constructs represent non-rational concepts), then the theory is a poor theory.
- **Explanatory power:** How much does a given theory explain (or predict) reality? Good theories obviously explain the target phenomenon better than rival theories, as often measured by variance explained (R-square) value in regression equations.
- **Falsifiability:** British philosopher Karl Popper stated in the 1940's that for theories to be valid, they must be falsifiable. Falsifiability ensures that the theory is potentially disprovable, if empirical data does not match with theoretical propositions, which allows for their empirical testing by researchers. In other words, theories cannot be theories unless they can be empirically testable. Tautological statements, such as "a day with high temperatures is a hot day" are not empirically testable because a hot day is defined (and measured) as a day with high temperatures, and hence, such statements cannot be viewed as a theoretical proposition. Falsifiability requires presence of rival explanations; it ensures that the constructs are adequately measurable, and so forth. However, note that saying that a theory is falsifiable is not the same as saying that a theory should be falsified. If a theory is indeed falsified based on empirical evidence, then it was probably a poor theory to begin with!
- **Parsimony:** Parsimony examines how much of a phenomenon is explained with how few variables. The concept is attributed to 14th century English logician Father William of Ockham (and hence called "Ockham's razor" or "Occam's razor"), which states that among competing explanations that sufficiently explain the observed evidence, the simplest theory (i.e., one that uses the smallest number of variables or makes the fewest

assumptions) is the best. Explanation of a complex social phenomenon can always be increased by adding more and more constructs. However, such approach defeats the purpose of having a theory, which are intended to be "simplified" and generalizable explanations of reality. Parsimony relates to the degrees of freedom in a given theory. Parsimonious theories have higher degrees of freedom, which allow them to be more easily generalized to other contexts, settings, and populations.

Approaches to Theorizing

How do researchers build theories? Steinfeld and Fulk (1990)² recommend four such approaches. The first approach is to build theories inductively based on observed patterns of events or behaviors. Such approach is often called "grounded theory building", because the theory is grounded in empirical observations. This technique is heavily dependent on the observational and interpretive abilities of the researcher, and the resulting theory may be subjective and non-confirmable. Furthermore, observing certain patterns of events will not necessarily make a theory, unless the researcher is able to provide consistent explanations for the observed patterns. We will discuss the grounded theory approach in a later chapter on qualitative research.

The second approach to theory building is to conduct a bottom-up conceptual analysis to identify different sets of predictors relevant to the phenomenon of interest using a predefined framework. One such framework may be a simple input-process-output framework, where the researcher may look for different categories of inputs, such as individual, organizational, and/or technological factors potentially related to the phenomenon of interest (the output), and describe the underlying processes that link these factors to the target phenomenon. This is also an inductive approach that relies heavily on the inductive abilities of the researcher, and interpretation may be biased by researcher's prior knowledge of the phenomenon being studied.

The third approach to theorizing is to extend or modify existing theories to explain a new context, such as by extending theories of individual learning to explain organizational learning. While making such an extension, certain concepts, propositions, and/or boundary conditions of the old theory may be retained and others modified to fit the new context. This deductive approach leverages the rich inventory of social science theories developed by prior theoreticians, and is an efficient way of building new theories by building on existing ones.

The fourth approach is to apply existing theories in entirely new contexts by drawing upon the structural similarities between the two contexts. This approach relies on reasoning by analogy, and is probably the most creative way of theorizing using a deductive approach. For instance, Markus (1987)³ used analogic similarities between a nuclear explosion and uncontrolled growth of networks or network-based businesses to propose a critical mass theory of network growth. Just as a nuclear explosion requires a critical mass of radioactive material to sustain a nuclear explosion, Markus suggested that a network requires a critical mass of users to sustain its growth, and without such critical mass, users may leave the network, causing an eventual demise of the network.

² Steinfeld, C.W. and Fulk, J. (1990). "The Theory Imperative," in *Organizations and Communications Technology*, J. Fulk and C. W. Steinfeld (eds.), Newbury Park, CA: Sage Publications.

³ Markus, M. L. (1987). "Toward a 'Critical Mass' Theory of Interactive Media: Universal Access, Interdependence, and Diffusion," *Communication Research* (14:5), 491-511.

Examples of Social Science Theories

In this section, we present brief overviews of a few illustrative theories from different social science disciplines. These theories explain different types of social behaviors, using a set of constructs, propositions, boundary conditions, assumptions, and underlying logic. Note that the following represents just a simplistic introduction to these theories; readers are advised to consult the original sources of these theories for more details and insights on each theory.

Agency Theory. Agency theory (also called principal-agent theory), a classic theory in the organizational economics literature, was originally proposed by Ross (1973)⁴ to explain two-party relationships (such as those between an employer and its employees, between organizational executives and shareholders, and between buyers and sellers) whose goals are not congruent with each other. The goal of agency theory is to specify optimal contracts and the conditions under which such contracts may help minimize the effect of goal incongruence. The core assumptions of this theory are that human beings are self-interested individuals, boundedly rational, and risk-averse, and the theory can be applied at the individual or organizational level.

The two parties in this theory are the principal and the agent; the principal employs the agent to perform certain tasks on its behalf. While the principal's goal is quick and effective completion of the assigned task, the agent's goal may be working at its own pace, avoiding risks, and seeking self-interest (such as personal pay) over corporate interests. Hence, the goal incongruence. Compounding the nature of the problem may be information asymmetry problems caused by the principal's inability to adequately observe the agent's behavior or accurately evaluate the agent's skill sets. Such asymmetry may lead to agency problems where the agent may not put forth the effort needed to get the task done (the *moral hazard* problem) or may misrepresent its expertise or skills to get the job but not perform as expected (the *adverse selection* problem). Typical contracts that are behavior-based, such as a monthly salary, cannot overcome these problems. Hence, agency theory recommends using outcome-based contracts, such as a commissions or a fee payable upon task completion, or mixed contracts that combine behavior-based and outcome-based incentives. An employee stock option plans are an example of an outcome-based contract while employee pay is a behavior-based contract. Agency theory also recommends tools that principals may employ to improve the efficacy of behavior-based contracts, such as investing in monitoring mechanisms (such as hiring supervisors) to counter the information asymmetry caused by moral hazard, designing renewable contracts contingent on agent's performance (performance assessment makes the contract partially outcome-based), or by improving the structure of the assigned task to make it more programmable and therefore more observable.

Theory of Planned Behavior. Postulated by Ajzen (1991)⁵, the theory of planned behavior (TPB) is a generalized theory of human behavior in the social psychology literature that can be used to study a wide range of individual behaviors. It presumes that individual behavior represents conscious reasoned choice, and is shaped by cognitive thinking and social pressures. The theory postulates that behaviors are based on one's intention regarding that behavior, which in turn is a function of the person's attitude toward the behavior, subjective

⁴ Ross, S. A. (1973). "The Economic Theory of Agency: The Principal's Problem," *American Economic Review* (63:2), 134-139.

⁵ Ajzen, I. (1991). "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* (50), 179-211.

norm regarding that behavior, and perception of control over that behavior (see Figure 4.2). Attitude is defined as the individual's overall positive or negative feelings about performing the behavior in question, which may be assessed as a summation of one's beliefs regarding the different consequences of that behavior, weighted by the desirability of those consequences. Subjective norm refers to one's perception of whether people important to that person expect the person to perform the intended behavior, and represented as a weighted combination of the expected norms of different referent groups such as friends, colleagues, or supervisors at work. Behavioral control is one's perception of internal or external controls constraining the behavior in question. Internal controls may include the person's ability to perform the intended behavior (self-efficacy), while external control refers to the availability of external resources needed to perform that behavior (facilitating conditions). TPB also suggests that sometimes people may intend to perform a given behavior but lack the resources needed to do so, and therefore suggests that posits that behavioral control can have a direct effect on behavior, in addition to the indirect effect mediated by intention.

TPB is an extension of an earlier theory called the theory of reasoned action, which included attitude and subjective norm as key drivers of intention, but not behavioral control. The latter construct was added by Ajzen in TPB to account for circumstances when people may have incomplete control over their own behaviors (such as not having high-speed Internet access for web surfing).

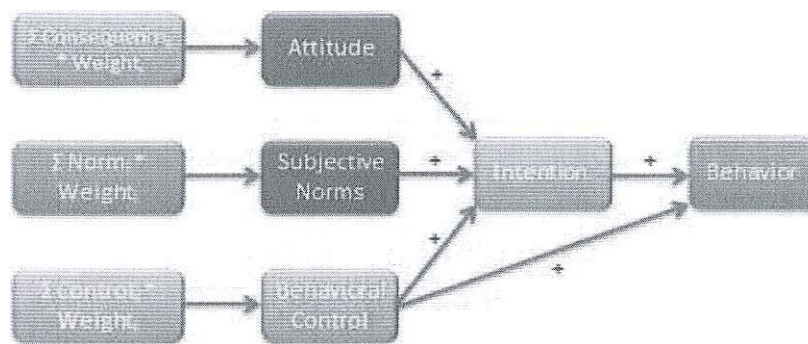


Figure 4.2. Theory of planned behavior

Innovation diffusion theory. Innovation diffusion theory (IDT) is a seminal theory in the communications literature that explains how innovations are adopted within a population of potential adopters. The concept was first studied by French sociologist Gabriel Tarde, but the theory was developed by Everett Rogers in 1962 based on observations of 508 diffusion studies. The four key elements in this theory are: innovation, communication channels, time, and social system. Innovations may include new technologies, new practices, or new ideas, and adopters may be individuals or organizations. At the macro (population) level, IDT views innovation diffusion as a process of communication where people in a social system learn about a new innovation and its potential benefits through communication channels (such as mass media or prior adopters) and are persuaded to adopt it. Diffusion is a temporal process; the diffusion process starts off slow among a few early adopters, then picks up speed as the innovation is adopted by the mainstream population, and finally slows down as the adopter population reaches saturation. The cumulative adoption pattern therefore forms an S-shaped curve, as shown in Figure 4.3, and the adopter distribution represents a normal distribution. All adopters are not identical, and adopters can be classified into innovators, early adopters, early majority, late majority, and laggards based on their time of their adoption. The rate of diffusion

also depends on characteristics of the social system such as the presence of opinion leaders (experts whose opinions are valued by others) and change agents (people who influence others' behaviors).

At the micro (adopter) level, Rogers (1995)⁶ suggests that innovation adoption is a process consisting of five stages: (1) knowledge: when adopters first learn about an innovation from mass-media or interpersonal channels, (2) persuasion: when they are persuaded by prior adopters to try the innovation, (3) decision: their decision to accept or reject the innovation, (4) implementation: their initial utilization of the innovation, and (5) confirmation: their decision to continue using it to its fullest potential (see Figure 4.4). Five innovation characteristics are presumed to shape adopters' innovation adoption decisions: (1) relative advantage: the expected benefits of an innovation relative to prior innovations, (2) compatibility: the extent to which the innovation fits with the adopter's work habits, beliefs, and values, (3) complexity: the extent to which the innovation is difficult to learn and use, (4) trialability: the extent to which the innovation can be tested on a trial basis, and (5) observability: the extent to which the results of using the innovation can be clearly observed. The last two characteristics have since been dropped from many innovation studies. Complexity is negatively correlated to innovation adoption, while the other four factors are positively correlated. Innovation adoption also depends on personal factors such as the adopter's risk-taking propensity, education level, cosmopolitanism, and communication influence. Early adopters are venturesome, well educated, and rely more on mass media for information about the innovation, while later adopters rely more on interpersonal sources (such as friends and family) as their primary source of information. IDT has been criticized for having a "pro-innovation bias," that is for presuming that all innovations are beneficial and will be eventually diffused across the entire population, and because it does not allow for inefficient innovations such as fads or fashions to die off quickly without being adopted by the entire population or being replaced by better innovations.

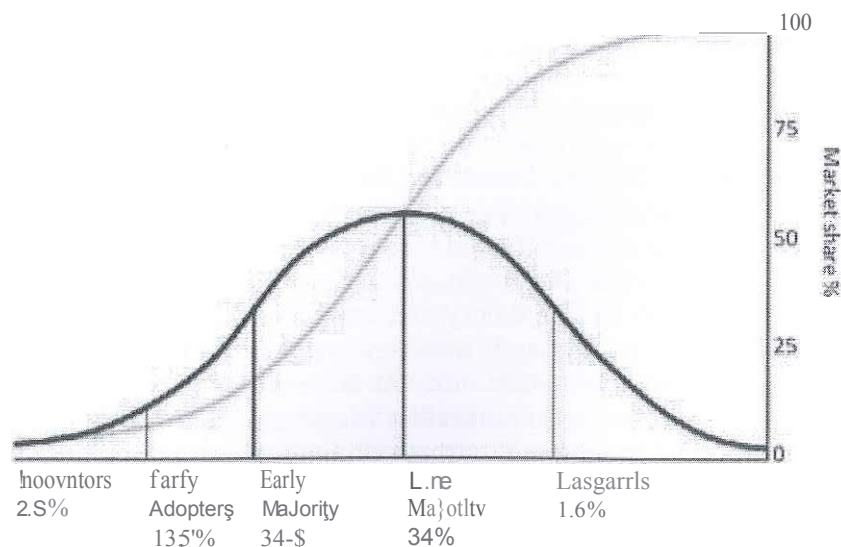


Figure 4.3. S-shaped diffusion curve

⁶ Rogers, E. (1962). *Diffusion of Innovations*. New York: The Free Press. Other editions 1983, 1996, 2005.

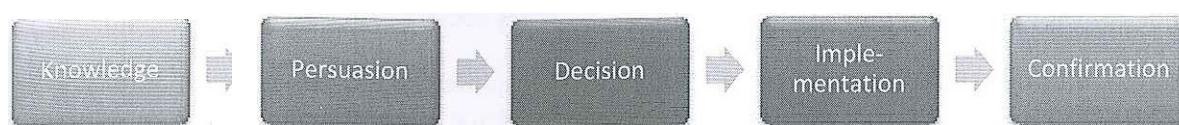


Figure 4.4. Innovation adoption process

Elaboration Likelihood Model. Developed by Petty and Cacioppo (1986)⁷, the elaboration likelihood model (ELM) is a dual-process theory of attitude formation or change in the psychology literature. It explains how individuals can be influenced to change their attitude toward a certain object, events, or behavior and the relative efficacy of such change strategies. The ELM posits that one's attitude may be shaped by two "routes" of influence, the central route and the peripheral route, which differ in the amount of thoughtful information processing or "elaboration" required of people (see Figure 4.5). The central route requires a person to think about issue-related arguments in an informational message and carefully scrutinize the merits and relevance of those arguments, before forming an informed judgment about the target object. In the peripheral route, subjects rely on external "cues" such as number of prior users, endorsements from experts, or likeability of the endorser, rather than on the quality of arguments, in framing their attitude towards the target object. The latter route is less cognitively demanding, and the routes of attitude change are typically operationalized in the ELM using the *argument quality* and *peripheral cues* constructs respectively.

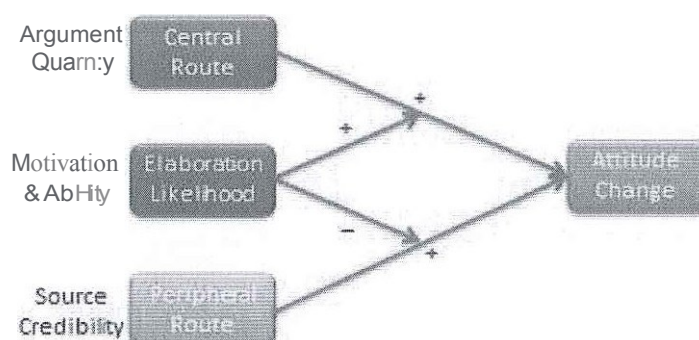


Figure 4.5. Elaboration likelihood model

Whether people will be influenced by the central or peripheral routes depends upon their ability and motivation to elaborate the central merits of an argument. This ability and motivation to elaborate is called *elaboration likelihood*. People in a state of high elaboration likelihood (high ability and high motivation) are more likely to thoughtfully process the information presented and are therefore more influenced by argument quality, while those in the low elaboration likelihood state are more motivated by peripheral cues. Elaboration likelihood is a situational characteristic and not a personal trait. For instance, a doctor may employ the central route for diagnosing and treating a medical ailment (by virtue of his or her expertise of the subject), but may rely on peripheral cues from auto mechanics to understand

⁷ Petty, R. E., and Cacioppo, J. T. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

the problems with his car. As such, the theory has widespread implications about how to enact attitude change toward new products or ideas and even social change.

General Deterrence Theory. Two utilitarian philosophers of the eighteenth century, Cesare Beccaria and Jeremy Bentham, formulated General Deterrence Theory (GDT) as both an explanation of crime and a method for reducing it. GDT examines why certain individuals engage in deviant, anti-social, or criminal behaviors. This theory holds that people are fundamentally rational (for both conforming and deviant behaviors), and that they freely choose deviant behaviors based on a rational cost-benefit calculation. Because people naturally choose utility-maximizing behaviors, deviant choices that engender personal gain or pleasure can be controlled by increasing the costs of such behaviors in the form of punishments (countermeasures) as well as increasing the probability of apprehension. Swift, severity, and certainty of punishments are the key constructs in GDT.

While classical positivist research in criminology seeks generalized causes of criminal behaviors, such as poverty, lack of education, psychological conditions, and recommends strategies to rehabilitate criminals, such as by providing them job training and medical treatment, GDT focuses on the criminal decision making process and situational factors that influence that process. Hence, a criminal's personal situation (such as his personal values, his affluence, and his need for money) and the environmental context (such as how protected is the target, how efficient is the local police, how likely are criminals to be apprehended) play key roles in this decision making process. The focus of GDT is not how to rehabilitate criminals and avert future criminal behaviors, but how to make criminal activities less attractive and therefore prevent crimes. To that end, "target hardening" such as installing deadbolts and building self-defense skills, legal deterrents such as eliminating parole for certain crimes, "three strikes law" (mandatory incarceration for three offenses, even if the offenses are minor and not worth imprisonment), and the death penalty, increasing the chances of apprehension using means such as neighborhood watch programs, special task forces on drugs or gang-related crimes, and increased police patrols, and educational programs such as highly visible notices such as "Trespassers will be prosecuted" are effective in preventing crimes. This theory has interesting implications not only for traditional crimes, but also for contemporary white-collar crimes such as insider trading, software piracy, and illegal sharing of music.

INTRODUCTION TO

SOCIAL RESEARCH

QUANTITATIVE & QUALITATIVE APPROACHES

KEITH F PUNCH

ISAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

(\$)**SAGE**

Los Angeles | London | New Delhi
Singapore | Washington DC

SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/1 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110044

SAGE Publications Asia-Pacific Pte Ltd
3 Church Street
#10-04 Samsung Hub
Singapore 049483

Editor: Katie Metzger
Development editor: Amy Jarrold
Production editor: Nicola Marshall
Copyeditor: Jennifer Hinchliffe
Proofreader: Sharon Cawood
Indexer: Silvia Benvenuto
Marketing manager: Ben Griffin-Sherwood
Cover designer: Francis Kenney
Typeset by: C&M Digital (P) Ltd, Chennai, India
Printed and bound in Great Britain by Ashford
Colour Press Ltd

© Keith F. Punch 2014
Chapter 3 © Alis Oancea 2014
Chapter 13 © Wayne McGowan 2014

First edition published 1998. Reprinted 1999, 2000, 2001, 2003, 2004.
Second edition published 2005. Reprinted 2007, 2008, 2009, 2010,
2011 (twice), 2012.
Third edition published 2014.

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.

All material on the accompanying website can be printed *off* and photocopied by the purchaser/user of the book. The web material itself may not be reproduced in its entirety for use by others without prior written permission from SAGE. The web material may not be distributed or sold separately from the book without the prior written permission of SAGE. Should anyone wish to use the materials from the website for conference purposes, they would require separate permission from us. All material © Keith F. Punch, 2014.

Library of Congress Control Number: 2013935252

British Library Cataloguing in Publication data

A catalogue record for this book is available from
the British Library

ISBN 978-1-4462-4092-2
ISBN 978-1-4462-4093-9 (pbk)

List of tables	xiii
List of figures	xv
Preface	xvii
Companion website	xix
About the author	xxi
1 Introduction	1
2 Theory and Method in Social Science Research	13
3 Ethics in Social Science Research	35
4 Research Questions	57
5 From Research Questions to Data	73
6 Literature Searching and Reviewing	93
7 Qualitative Research Design	113
8 Collecting Qualitative Data	143
9 Analysing Qualitative Data	167
10 Quantitative Research Design	205
11 Collecting Quantitative Data	227
12 Analysing Quantitative Data	251
13 The Internet and Research	279
14 Mixed Methods and Evaluation	301
15 Research Writing	327
Glossary	347
Appendix 1	355
References	359
Index	377

two

THEORY AND METHOD SOCIAL SCIENCE RESEARCH

Contents

2.1 Methodological theory

2.2 Substantive theory

Description versus explanation

Theory verification-theory generation

Question-method connections

Prespecified versus unfolding: structure in research questions, design and data

Chapter summary

Key terms

Exercises and study questions

Further reading

Notes

–Learning Objectives–

After studying this chapter you should be able to:

- Describe what is meant by methodological theory and by substantive theory
- Define paradigms, and describe the difference between paradigm-driven research and pragmatic research
- Understand the difference between description and explanation
- Describe the difference between theory verification and theory generation research
- Explain the logical priority of research questions over research methods
- Describe the essential differences between prespecified and unfolding research

The term 'theory' is used in many different ways in the literature, which can create difficulties. In this chapter, I focus on two main uses of theory – methodological theory and substantive theory. Both are important. Methodological theory concerns the theory or philosophy behind research methods, and is discussed in Section 2.1. It leads on to the topic of question-method connections (Section 2.5). Substantive theory concerns the content area of research, and is discussed in Section 2.2. It leads on to the topics of description and explanation (Section 2.3), and to theory verification and theory generation (Section 2.4). The final section of the chapter deals with the issue of structure in planning a piece of research.

Methodological theory

Methodological theory, as used here, means theory about method. Whereas substantive theory is about substance or content, methodological theory is about method – about what lies behind the approaches and methods of inquiry used in a piece of research.

Methods of inquiry are based on assumptions – assumptions about the nature of the reality being studied, assumptions about what constitutes knowledge of this reality, and assumptions about what therefore are appropriate methods of building knowledge of this reality. Very often these assumptions are implicit. A point of contention in research methods training has often been whether or not it should be required that such assumptions are made explicit in a piece of postgraduate research.

These assumptions constitute the essential idea of what is meant by the term 'paradigm' in the research methodology and philosophy of science literature. Paradigm issues are necessarily philosophical in nature. In general, paradigm means a set of assumptions about the world, and about what constitute proper topics and techniques for inquiring into that world. Put simply, it is a way of looking at the world. It means a view of how inquiry should be done (hence the term 'inquiry paradigm' which is sometimes used), and is a broad term encompassing elements of epistemology, theory and philosophy, along with methods.

Denzin and Lincoln (1994: 107-9) describe a paradigm as:

a set of basic beliefs (or metaphysics) that deals with ultimates or first principles. It represents a worldview that defines, for its holder, the nature of 'the world,' the individual's place in it, and the range of possible relationships to that world and its parts.

They point out that inquiry paradigms define what they are concerned with, and what falls within and outside the limits of legitimate inquiry, and that inquiry paradigms address three fundamental questions, which reflect the assumptions noted above:

• The ontological question: What is the form and nature of reality and, therefore, what is there that can be known about it?

The epistemological question: What is the relationship between the knower and what can be known?

The methodological question: How can the inquirer go about finding out what can be known?

In simpler language, paradigms tell us:

what the reality is like (ontology);

what the relationship is between the researcher and that reality (epistemology); and

what methods can be used for studying the reality (methodology).

These three interrelated questions illustrate the connections between methods and the deeper underlying philosophical issues. Methods are ultimately based on, and derive from, paradigms. Conversely, paradigms have implications for methods. This point became clear during methodological developments of the past 40-50 years. At this point, therefore, a brief sketch of some historical background on methods and paradigms in social science research is appropriate.

Beginning in the 1960s, the traditional dominance of quantitative methods, as the way of doing empirical social science research, was challenged. This challenge accompanied a major growth of interest in using qualitative methods, and this in turn produced a split in the field, between quantitative and qualitative researchers. A prolonged quantitative-qualitative debate ensued, sometimes described as the 'paradigm wars'.¹

Much of that debate was characterised by either/or thinking. Some thought that only quantitative approaches should be used in research. Others were just as emphatic that only qualitative approaches are appropriate. More recently, however, there have been moves towards a detente, and an increased interest in the combination of the two approaches (Bryman, 1988, 1992; Hammersley, 1992; Tashakkori and Teddlie, 2003a). This has led to mixed methods, the topic of Chapter 14, and a major growth area in the recent research methodology literature. These methodological changes have occurred across most areas of empirical social science research, though in some areas the changes have been more pronounced than in others.

The full story of these developments and debates is more complex than this. I have focused only on one main dimension of it, the quantitative-qualitative distinction, because these remain two of the central methodological approaches in social science research today, and because this distinction is a central organising principle for this book. A major consequence of these developments is that qualitative research methods have moved much more into the mainstream of social science research, compared with their marginalised position of 40 or so years ago. As noted, a further development has been the combination of the two approaches in what is now called 'mixed methods research' (see Chapter 14). As a result, the field of research methodology in social science is now bigger and more complex than it used to be.

Because of the connections between methods and paradigms, the history briefly outlined above also has a deeper level, a level that is not just about the quantitative-qualitative debate, or about research methods, but about paradigms themselves. On this deeper level, a major rethinking began some time ago, and is ongoing. It has brought a questioning of all aspects of research (its purposes, its place and role, its context and conceptualisations of research itself) as well as the methods it uses. It has also brought the development of new perspectives, and of new approaches to data and to the analysis of data, within qualitative research especially. Prominent features of this rethinking are the detailed critique of positivism, and the emergence and articulation of several different paradigms, as alternatives to positivism. As a result, paradigm issues are in a state of change and development, and many matters are still contested.

It is the development of qualitative methods which has exposed the many different paradigm possibilities, and the situation has now become very complicated. Thus Denzin and Lincoln (1994: 109) identify four main alternative inquiry paradigms underlying qualitative research (positivism, post-positivism, critical theory, constructivism), but more detailed examples and classifications of paradigms are given by Guba and Lincoln (1994). Morse (1994: 224-5) has this classification of paradigms with associated qualitative research strategies: philosophy-phenomenology; anthropology-ethnography; sociology-symbolic interactionism-grounded theory; semiotics-ethnomethodology and discourse analysis. Janesick (1994: 212) has a more detailed list of paradigm-related qualitative research strategies, noting that it is not meant to include all possibilities: ethnography, life history, oral history, ethnomethodology, case study, participant observation, field research or field study, naturalistic study, phenomenological study, ecological descriptive study, descriptive study, symbolic interactionist study, microethnography, interpretive research, action research, narrative research, historiography and literary criticism. And examples of paradigms considered by writers in the philosophy of education are logical empiricism and post-empiricism, critical rationalism, critical theory, phenomenology, hermeneutics and systems theory.

This can be confusing and daunting territory for the beginning researcher, partly because of philosophy and partly because of terminology. Fortunately, in the light of these complications, some of the literature now seems to be converging and simplifying. In one version of this convergence, the main paradigm positions are

positivism and interpretivism; in another they are positivism and constructivism. Thus we have:

positivism (associated mostly with quantitative methods), and
either interpretivism or constructivism (associated with qualitative methods).

These associations – positivism with quantitative methods and interpretivism-constructivism with qualitative methods – are generally true, but they are not necessary associations. It is more accurate to say that positivism is likely to be associated with quantitative methods, and interpretivism and constructivism are likely to be associated with qualitative methods.

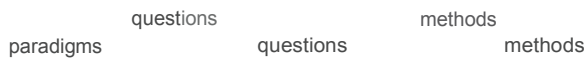
These terms are defined slightly differently by different writers, but their main nature-of-reality ideas are as follows:

Positivism – the belief that objective accounts of the world can be given, and that the function of science is to develop descriptions and explanations in the form of universal laws – that is, to develop nomothetic knowledge.

Interpretivism – concentrates on the meanings people bring to situations and behaviour, and which they use to make sense of their world (O'Donoghue, 2007: 16-17); these meanings are essential to understanding behaviour.

Constructivism – realities are local, specific and constructed; they are socially and experientially based, and depend on the individuals or groups holding them (Guba and Lincoln, 1994: 109-11).

In Section 2.5, question-method connections are discussed, and I stress that there needs to be compatibility and integrity in the way the research questions and research methods fit together in a study. This is shown in the top line in the diagram below. Paradigms expand that, because paradigms have implications both for the sorts of research questions asked and the methods used to answer them. This is shown in the bottom line in the diagram.



What does all this methodological theory mean for planning and executing a piece of research? Broadly, there are two main ways in which planning a research project can proceed:

Paradigm-driven approach – one way is to begin with a paradigm, articulate it and develop research questions and methods from it;

Pragmatic approach – the other way is to begin with research questions that need answers and then choose methods for answering them.

In the pragmatic approach, the questions may come from any source – the literature, existing substantive theory, the media, personal experience, and so on. But very often, especially in professional fields such as education, management or nursing, they will come from practical and professional issues and problems associated with

the workplace. The starting point here is not a paradigm. Instead, the starting point is a problem that needs a solution or a question that needs answers. This is a pragmatic approach.

This has sometimes been a contentious issue in higher-degree research programmes. Some university departments have taken the view that paradigm issues are paramount, and insist that research should not be allowed to proceed until it has articulated its paradigm position. I believe this insistence is not well placed, because paradigm-driven research is not the only way to proceed, and because I see a big role for a more pragmatic, applied and professional approach to social science research. I have no objection to paradigm-driven research. My objection is only to the view that all research must be paradigm-driven. I take a similar view with respect to the philosophical issues involved in paradigm debates. I think we should be aware of the issues involved, and of the areas of debate. These are indicated in several places throughout the book. But we can proceed to do research, and to train researchers, mindful of those debates yet not engulfed by them, and without necessarily yet being able to see their resolution. In other words, we can acknowledge the connections of methods to these deeper issues, and discuss them from time to time as they arise, without making them the major focus of our research. This is to take the pragmatic approach noted, consistent with the view that not all questions for social research are driven by paradigm considerations, and that different sorts of questions require different methods for answering them. Both of these points are elaborated upon in later chapters.

To choose the pragmatic approach is to start by focusing on what we are trying to find out in research, and then to fit methods in with this. The important topic of question-method connections is discussed in Section 2.5.

Contentive theory

By substantive theory I mean theory about a substantive issue or phenomenon, some examples of which are shown below. Substantive theory is content-based theory, and is not concerned with methods. Its purpose is to explain some phenomenon or issue of interest – it is explanatory theory. But because explanation requires description (see Section 2.3), substantive theory both describes and explains. An explanatory theory both describes and explains the phenomenon of substantive interest. Theory, in this sense, is a set of propositions that together describe and explain the phenomenon being studied. These propositions are at a higher level of abstraction than the specific facts and empirical generalisations (the data) about the phenomenon. They explain the data by deduction, in the if-then sense. This is the model of scientific knowledge shown in Figure 2.1.

Some examples of substantive theories from different areas of social research are attribution theory, reinforcement theory, various learning theories and personal construct theory (from psychology); reference group theory and social stratification theory (from sociology); the theory of vocational personalities and career anchors (from occupational sociology); various leadership theories (from management and

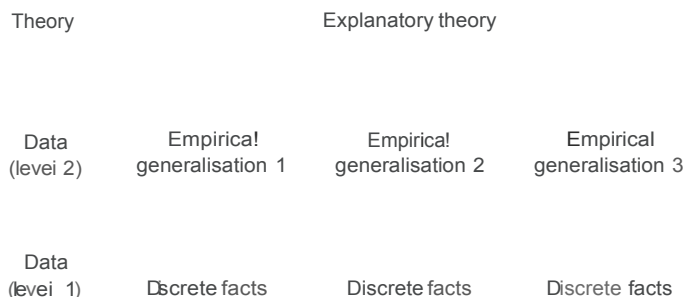


FIGURE 2.1 The relationship between theory and data

administration), and theories of children's moral development and of teacher career cycles (from education).

Thus an important question in planning research is 'What is the role of (substantive) theory in this study?' This question is sometimes considered more appropriate for doctoral-level research than for masters-level research. This seems to be because a common criterion among universities for the award of the doctorate centres on the 'substantial and original contribution to knowledge' a study makes, and the 'substantial' part of that criterion is often interpreted in terms of its contribution to substantive theory.

Description versus explanation

In Chapter 1 a brief description of the scientific method was given, stressing that it has the two central parts of data and theory, and that the objective of scientific inquiry is to build explanatory theory about its data. In this view, the aim is to explain the data, not just to use the data for description. This distinction between description and explanation is particularly relevant to the purposes of a piece of research.

The description-explanation distinction is easy to understand on one level, and difficult to understand on another.² Fortunately, it is on the easier level where the practical value of the distinction lies. Description and explanation represent two different levels of understanding. To describe is to somehow draw a picture of what happened, or of how things are proceeding, or of what a situation or person or event is like. To explain, on the other hand, is to account for what happened, or for how things are proceeding, or for what something or someone is like. It involves finding the reasons for things (or events or situations), showing why and how they have come to be what they are. Description is a more restricted purpose than explanation. We can describe without explaining, but we cannot really explain without describing. Therefore explanation goes further than description. It is more than just description – it is description plus something else.

Description focuses on what is the case, whereas explanation focuses on why (and sometimes how) something is the case. Science as a method of building knowledge has, in general, pursued the objective of explanation, not just of description. There is a good reason for *this*. When we know why something happens, we know much more than just what happens. It puts us in a position to predict what will happen, and perhaps to be able to control what will happen.

Thus explanatory knowledge is more powerful than descriptive knowledge. But descriptive knowledge is still important, since explanation requires description. To put it around the other way, description is a first step towards explanation. If we want to know why something happens, it is important to have a good description of exactly what happens. There are often clues to explanation in a full description, and it is hard to explain something satisfactorily until you understand just what the something is (Miles, Huberman and Saldana, 2013).

This distinction comes up mainly when the purpose of a piece of research is being considered. Is the purpose to describe, to explain or both? Descriptive studies are sometimes given a lower status than studies that aim to explain. That is why we sometimes hear the expression 'it is only a descriptive study'. But while this judgement may sometimes have merit, it has to be made carefully. There are situations where a thorough descriptive study will be very valuable. Two examples of such situations are:

when a new area for research is being developed, and initial and exploratory studies are planned – it is very sensible then to focus on systematic description as the objective of the research;

careful description of complex social processes can help us to understand what factors to concentrate on for later explanatory studies.

Whether description or explanation is the appropriate purpose for a piece of research depends on the particular situation. Here, as elsewhere, blanket rules are not appropriate. Rather, each research situation needs to be analysed and understood in its own context. It is useful to raise this question of whether the objective of a study is description and/or explanation, especially during the planning stages of research. A good way to do it is to ask 'why' about the things being studied, as well as 'what'.

Thus explanation is the central focus of substantive theory. The essential idea is to explain what is being studied, with the explanation being couched in more abstract terms than the terms used to describe it.³ We will return to this idea of theory in two places later in the book. The first is in Chapter 4 (Section 4.7), where we consider the role of hypotheses in relation to research questions. There we will see that theory stands behind the hypothesis, in an inductive-deductive relationship with it (Brodbeck, 1968; Nagel, 1979). Studies that use this approach are theory verification studies. The second is in Chapter 9, where we discuss grounded theory analysis in studies that aim to develop theory. These are theory generation studies.

Theory verification-theory generation

This distinction between theory verification and theory generation research is important. A project that has explanation as its objective can set out to test theory, or to build theory – to verify theory, or to generate it. For Wolcott (1992), this is the distinction between 'theory first' and 'theory after'. In theory-first research, we start with a theory, deduce hypotheses from it and design a study to test these hypotheses. This is theory verification. In theory-after research, we do not start with a theory. Instead, the aim is to end up with a theory, developed systematically from the data we have collected. This is theory generation.

Quantitative research has typically been more directed at theory verification, while qualitative research has typically been more concerned with theory generation. While this correlation is historically valid, there is no necessary connection between purpose and approach. That is, quantitative research can be used for theory generation (as well as for verification) and qualitative research can be used for theory verification (as well as for generation), as pointed out by various writers (for example, Hammersley, 1992; Brewer and Hunter, 2005). However, while the connection is not necessary, it is nonetheless likely that theory generation research will more often be qualitative. Research directed at theory generation is more likely when a new area is being studied, and exploration of this new area is more likely to use the less structured fieldwork techniques of qualitative research.

Is theory verification research better than theory generation research? This book does not favour one research purpose over the other, since both are needed and both have their place. Either purpose can be appropriate in a research project, and sometimes both will be appropriate. It depends on the topic, the context and practical circumstances of the research, and especially on how much prior theorising and knowledge exists in the area. As with other aspects of a project, the researcher needs to consider the alternatives, select among them according to consistent and logical criteria, and then articulate that position.

Theory generation research was given new legitimacy in social science by the development of grounded theory. As is described in Chapter 7, grounded theory is an explicit theory generation research strategy, developed in reaction against the overemphasis on theory verification research in the American sociology of the 1940s and 1950s. Glaser and Strauss stated this clearly in their original grounded theory publication:

Verification is the keynote of current sociology. Some three decades ago, it was felt that we had plenty of theories but few confirmations of them – a position made very feasible by the greatly increased sophistication of quantitative methods. As this shift in emphasis took hold, the discovery of new theories became slighted and, at some universities, virtually neglected. (Glaser and Strauss, 1967: 10)

Glaser and Strauss argued that the emphasis on verification of existing theories kept researchers from investigating new problem areas, prevented them from

acknowledging the necessarily exploratory nature of much of their work, encouraged instead the inappropriate use of verificational logic and rhetoric, and discouraged the development and use of systematic empirical procedures for generating as well as testing theories (Brewer and Hunter, 2005).

This gives us a useful general guideline for when each purpose might be appropriate. When an area has lots of unverified theories, an emphasis on theory verification research seems a good thing. On the other hand, when an area is lacking in appropriate theories, it is time for the emphasis to shift to theory generation. Also, when research is directed mostly at the verification of existing theories, looking at new problem areas is discouraged, and the logic and techniques (usually quantitative) of verification research are seen as more important. When it is important to look at new areas in research, theory generation appeals as the appropriate purpose. This aspect of grounded theory research is taken up again in Chapter 7 (Section 7.5).

The description-explanation distinction fits in with the structure of scientific knowledge shown in Figure 2.1. In line with the conception of science given in Chapter 1, we can distinguish three levels of knowledge. At the lowest level, there are discrete facts. At the next level are empirical generalisations which group those facts together. At the highest level are theories, whose function is to explain the generalisations. This structure is summarised in the diagram shown. The first two levels (facts and empirical generalisations) focus on description, while the third level focuses on explanation.

This model of the structure of scientific knowledge comes primarily from a positivistic perspective, and stresses a nomothetic view of knowledge. It can be contrasted with an ideographic view of knowledge, a more appropriate aspiration for research in the eyes of many qualitative researchers.⁴ But while acknowledging its nomothetic bias, this model is very useful as a starting point in learning about social science research. Much research is based on this model, and it can often help in organising an individual project. It is clear and easy to understand, so the researcher who wishes to diverge from this model can see where and why the divergence occurs. In other words, when researchers argue about how research should proceed and contribute to knowledge, this model helps to see what the argument is about.

There is another reason for stressing this model here. It shows the hierarchical structure of knowledge, with higher levels of abstraction and generality at the top and lower levels at the bottom. This is similar to the hierarchical structure that links data indicators to variables and concepts, and which is central both to the concept-indicator model behind grounded theory coding in qualitative research, and to latent trait measurement theory in quantitative research. These topics are described in Chapters 9 and 11 respectively. This hierarchical structure of increasing levels of abstraction and generality, shown here with respect to scientific knowledge in general, and shown in later chapters with respect to concept-data links in both quantitative and qualitative research, is thus fundamental to much empirical research. An illustration of it is given in Example 2.1.

The Hierarchical Structure of Knowledge

A classic example of this way of structuring knowledge is Durkheim's work on the social aetiology of suicide, described in Durkheim (1951) and summarised in Greenwood (1968). Durkheim theorises 'upwards' from a series of empirical generalisations to a Law of suicide.⁵

Question-method connections

The principle here is that the matching or fit between the research questions and research methods should be as close as possible. A very good way to do that is for methods to follow from questions.

Different questions require different methods to answer them. The way a question is asked has implications for what needs to be done, in research, in order to answer it. Quantitative questions require quantitative methods to answer them, and qualitative questions require qualitative methods to answer them. In today's research environment, with quantitative and qualitative methods often used alongside each other, the matching of questions and methods is even more important. Since this book deals directly with both approaches, it is inevitable that this issue should be a recurrent concern.

The wording of questions is also important, since some wordings carry methodological implications. Thus research questions that include such terms as 'variables', 'factors that affect' and 'the determinants or correlates of', for example, imply a quantitative approach, while questions that include such terms as 'discover', 'seek to understand', 'explore a process' and 'describe the experiences' imply a qualitative approach. (Creswell, 2013 links these last four terms to grounded theory, ethnography, case study and phenomenology respectively).

An example of different research questions and their implications for methods is given by Shulman, in education research (1988: 6-9). He takes the study of reading, suggests four different types of questions, and shows the methods that would be required to answer each.

- 1 A first question might be: What makes some people successful readers and others unsuccessful? (Or, how can we predict what sorts of people will have difficulty learning to read?) Such questions would be answered using a quantitative correlational study that examined relationships between variables.
- A second question might be: What are the best possible methods for teaching reading to youngsters, irrespective of their backgrounds or attitudes? This question would involve a quantitative experimental study comparing different teaching methods.
- A third question might be: What is the general level of reading performance across different age, sex, social or ethnic groups in the population? This would require a quantitative survey of reading performance and reading practices.

A fourth set of questions might be quite different from the previous ones: How is reading instruction carried on? What are the experiences and perceptions of teachers and students as they engage in the teaching and learning of reading? How is this complex activity accomplished? Here, a qualitative case study involving observation and interview might be used, perhaps using the perspective of ethnomethodology.

Shulman goes on to suggest philosophical and historical questions as well. Other illustrations of question-method connections are given in Example 2.2.

EXAMPLE 2.2

Question-Method Connections

Shulman (1988: 6-9) shows connections between questions and methods with the topic of reading research in education; similar examples are noted by Seidman (2013). Marshall and Rossman (2010) show, in a table, the links between research purposes, research questions, research strategy and data collection techniques.

Maxwell (2012) adapts a table from LeCompte and Preissle (1993) to show the links between 'What do I need to know?' and 'What kind of data will answer the questions?' and illustrates these links with actual research questions.

Maxwell (2012) gives the example of a mismatch between questions and method, whereby, in a study of how historians work, the 'right answer' is found to be to the 'wrong question'.

A good way to achieve a fit between questions and methods is to ensure that the methods we use follow from the questions we seek to answer. In other words, the content of the research (the research questions) has a logical priority over the method of the research. To say that content precedes method is simply to say that we first need to establish what we are trying to find out, and then consider how we are going to do it. On a practical level, this is often a good way to get a research project off the ground. Sometimes it is difficult to know where and how to start, in planning research. If so, asking 'What are we trying to find out?' usually gets our thinking going, and ensures that we start with the content, not with the method. Putting questions before methods is also a good defence against overload when developing a research proposal. To delay consideration of methods until it is clear what the questions are helps in managing the inevitable complications that accompany a full examination of the possibilities for research in any area. It helps in keeping the question development stage systematic, and under control. It also helps achieve good question-method fit, a central criterion in the validity of research.

I am stressing this point here to counter a previous unfortunate tendency in social science research. In Chapter 1, the term 'methodolatry' was used:

I use the term *methodolatry*, a combination of *method* and *idolatry*, to describe a pre-occupation with selecting and defending methods to the exclusion of the actual substance of the story being told. Methodolatry is the slavish attachment and devotion to method that so often overtakes the discourse in the education and human service fields. (Janesick, 1994: 215)

Methodolatry means putting method before content. It is first learning the research method, then finding research questions that can fit into the method. It is looking for research questions guided by methods.

This is a danger when we place too much stress on the teaching of research methods, for their own sake. Because of this danger, this book concentrates on the logic and rationale behind empirical research and its methods. Once this logic is mastered, we can focus on research questions, and then fit the techniques and methods to the questions. In my opinion, the best sequence of learning activities for research is to start by learning the logic of research, then to focus on identifying and developing the research questions, and then to fit methods and techniques to the questions.

I am using the concept of methodolatry to argue for minimising the direct influence of methods on research questions, which we can do by first getting the research questions clear, and then focusing on the methods required to answer them. But methods can also indirectly influence research questions, by constraining what can be studied. There are limits as to what can be designed in research, and to what data can be obtained and analysed. While taking this into account, the advice is nonetheless to focus on questions first, as much as possible. In the above example, after showing how different methodological approaches fit different questions, Shulman emphasises the same point: 'we are advised to focus first on our problem and its characteristics before we rush to select the appropriate method' (1988: 15). Thus, when misfit between the parts becomes apparent during the planning of the research, it is a matter of adapting the parts to each other.

Question-method fit is an aspect of conceptual clarity in a piece of research. Conceptual clarity involves the precise and consistent use of terms, internal consistency within an argument and logical links between concepts, especially across different levels of abstraction. The pre-empirical question development work described in Chapter 4 is directed at this conceptual clarity. Developing specific research questions is a good way of achieving clarity and matching questions and methods.

The different paradigms and strategies within qualitative research open up many new and different types of research questions. For example, ethnographic questions might focus on cultural and symbolic aspects of behaviour; grounded theory questions might focus on understanding social processes, and how people manage different types of situations; a conversation analysis study might focus on conversational structure and on the role of conversation in taken-for-granted everyday activities; discourse analysis questions might focus on the way an institution presents itself to the world, the symbols and language it uses, and the connection of those with its ideology, knowledge, power, and so on. Paradigms can thus be important in

generating research questions. Within qualitative research especially, the range of questions of interest is now very broad. But it remains important, even with this broader range of questions, that the methods we use should follow from and fit in with the questions we seek to answer.

Prespecified versus unfolding: structure research questions design and data

How much should the research questions, design and data be preplanned in a piece of research, and how much should they emerge (or unfold) as the research develops?

There is a continuum we can set up for thinking about this question, with the dimension of interest being the amount of prespecified structure in the research strategy that is used. The central comparison is between research that is prespecified (or preplanned, or prefigured, or predetermined) on the one hand, and research that is unfolding (or emerging, or open-ended) on the other. Prespecified here refers to how much structure is introduced ahead of the empirical work, as opposed to during the empirical work. This continuum applies to three main areas – to research questions, to research design and to data.

Miles, Huberman and Saldana (2013) discuss this idea in the context of qualitative research under the heading of 'tight versus loose'. Those terms are equivalent to the terms used here – tight means prespecified and loose means unfolding. The key questions are: To what extent are the research questions, the design and the data focused, specified and structured ahead of the actual empirical work? To what extent does the focus in the research questions, and the structure in the design and the data, unfold and emerge as the empirical work proceeds? The continuum of possibilities is shown in Figure 2.2. This diagram shows that quantitative research typically falls towards the left-hand end of the continuum, whereas qualitative research can occupy a much greater range along the continuum.

'Structure', as used here, means showing what the different parts of the research are, how they connect with each other, what will be done in the research, and in what sequence. It means knowing what we are looking for, and how we are going to



FIGURE 2.2 Prespecified versus unfolding: the timing of structure

et it _ knowing what data we will want, and how they will be collected. It also means knowing what structure the data will have and how they will be analysed.

At the extreme left-hand end of the continuum, everything is prespecified – the research questions, the design and the data. It is all worked out in advance, a set of steps is laid down, and the researcher proceeds through those steps. At the other end, we can envisage a project where very little structure is determined in advance, with an open-ended and unstructured approach to the research questions, the design and the data. The strategy is that these will unfold as the study is carried out. Let us see what this contrast means for each of the three areas.

Research questions: at the left-hand end of the continuum, specific research questions are set up in advance to guide the study. It is quite clear, in advance, exactly what questions the study is trying to answer. At the right-hand end, only general questions are set up in advance. The argument there is that, until some empirical work is carried out, it is not possible (or, if possible, not sensible) to identify the specific research questions. They will only become clear as the research unfolds, and as a specific focus for the work is developed. Wolcott (1982) describes this contrast as 'looking for answers' versus 'looking for questions'. As we will see in Chapter 5, there is often a close connection between the research questions and the conceptual framework in a study. The issue described here in terms of research questions applies to conceptual frameworks as well – they can be developed and specified in advance of the research, or they can emerge as the research proceeds. The more tightly developed and pre-specified the research questions are, the more likely it is that there will be a well-developed conceptual framework as well.

Design: at the left-hand end, the design is tightly structured. The clearest examples come from quantitative research – experimental studies and non-experimental quantitative studies with carefully developed conceptual frameworks. Research questions, design and conceptual framework come together here, since a tightly structured design requires that variables be identified, and that their conceptual status in the research be made clear. At the right-hand end, the design is indicated in general terms only (for example, as in an unfolding case study, or an ethnography). Like the research questions, it will take detailed shape as the research progresses, and as the specific focus for the study is developed.

Data: at the left-hand end, data are structured in advance. A very clear example is quantitative data, where measurement is used to give the data numerical structure. Using numbers is the most common way of structuring data in advance, but there are other ways as well. Whether it is numerical or other categories, the point is that those categories are pre-established, or set up *a priori*. At the right-hand end, the data are unstructured at the point of collection. No pre-established categories or codes are used. The structure of the data, the categories and codes, emerge from the data, during the analysis – they are developed *a posteriori*. Thus the comparison is between starting with categories for the data, versus getting to them during the analysis of the data – between pre-coding the data and post-coding the data. This point about data has implications for instrumentation in data collection, not only in quantitative research, but in qualitative research as well.

The continuum shown in Figure 2.2 can now be described more accurately. It is really about when in the research process the structure is introduced. The structure can be introduced in the planning or pre-empirical stage, when the research is being

set up, before data are collected; or, it can be introduced in the execution stage of the research, as the study is being carried out, as data are being collected. Either way, structure is needed. A research project will be difficult both to report and to understand, and will lack credibility as a piece of research, without structure in its research questions, its design, especially in its data, and also in its report. So this contrast is not about having structure or not having structure, but about when in the research process the structure occurs. In other words, this continuum is about the timing of structure in the research – whether that structure is introduced ahead of the empirical research, or is introduced during and as a result of the empirical research.

The possibilities along this continuum represent different possible research styles. As the diagram shows, there is a correlation between these styles, on the one hand, and the typical quantitative and qualitative research approaches on the other. The typical quantitative study is much more likely to have specific research questions, a clear conceptual framework and design for its variables, and to use measurement as its way of structuring the data. It is harder to talk about typical qualitative studies, and they may cover a wider range along the continuum. Many of them fall towards the right-hand end, with general rather than specific questions set up in advance, with only a general design and with data not coded at the point of collection. This is well captured by Miles and Huberman (1994: 17), in discussing field research as a central part of the qualitative approach:

The conventional image of field research is one that keeps prestructured designs to a minimum. Many social anthropologists and social phenomenologists consider social processes to be too complex, too relative, too elusive, or too exotic to be approached with explicit conceptual frames or standard instruments. They prefer a more loosely structured, emergent, inductively 'grounded' approach to gathering data: The conceptual framework should emerge from the field in the course of the study; the important research questions will become clear only gradually; meaningful settings and actors cannot be selected prior to fieldwork; instruments, if any, should be derived from the properties of the setting and its actors' views of them.

This general correlation between style and approach also extends to theory verification versus theory generation research, the distinction discussed in Section 2.4. Theory verification research, by definition, is more likely to have clear-cut research questions leading to hypotheses, a tightly structured design and pre-established categories for data. Theory generation research, by contrast, will more likely use an approach where specific research questions unfold as the study develops, and where codes and categories for the data are empirically derived.

It is not a question of which strategy is best, since a large part of the answer to this question is 'it depends'. The question interacts with the overall approach to the research. Is it a quantitative study, a qualitative study or one that combines the two approaches? If quantitative, it is more likely to be towards the left-hand end of the continuum in Figure 2.2. If qualitative, there is likely to be a greater range

of possibilities. Nor is it a dichotomous choice between two extreme positions – it is a continuum. For clarity, the description in this section has been given in terms of the ends of the continuum. In reality, there are many points along the continuum, and any study may combine elements of either strategy – the prespecified one or the unfolding one.

How much predetermined structure is desirable in a project is a matter for analysis in each particular research situation. Structure is necessary. But the timing of the structure – when is the appropriate point to introduce this structure – depends on such factors as the topics and goals of the research, the availability of relevant knowledge and theory about the topic, and the researcher's familiarity with the situation being studied (Miles, Huberman and Saldana, 2013). Other factors to be considered are the preferred style of the research, the resources (including time) available to the researcher, and to what extent the researcher is interested in explanation versus interpretation. Depending on these factors, there can be merit in either approach. As Miles, Huberman and Saldana (2013) point out, what is required is a careful analysis of each situation where research is proposed. The research strategy should then be custom-built, as far as is possible, on the basis of this analysis.

The discussion in this section has treated research questions, design and data together. Subsequent chapters deal with questions, design and data separately, before bringing them back together in Chapters 14 and 15. Without wishing to advise against exploratory unfolding studies, it is worth noting some of the benefits *in* having at least a reasonable level of specificity in the research questions. For example, they give guidance during initial data collection, thereby saving time and resources and helping to avoid confusion and overload, an especially valuable benefit for the beginning researcher. In addition, research questions that are at least reasonably focused make it easier to communicate about the research, which can be important in the presentation (and approval) of a research proposal. Brewer and Hunter (2005) point out that, once a study is completed, it is irrelevant whether the research questions initiated the study or emerged from it – but it can matter at the proposal stage. Finally, it is very often the case that the researcher does have knowledge about the proposed research problems, even in a relatively unexplored area ('experiential data' and 'experiential knowledge' – see Strauss, 1987 and Maxwell, 2012). There is great benefit in getting that knowledge out onto the table, and working carefully to develop research questions in advance of the empirical work is a good way to do that.

Developing specific research questions to a point where they are stable, and connecting them to the design, data collection and data analysis parts of the research, requires careful work. The question being considered here is whether that work *is* done in advance of the research or during it. That brings us back to fitting the various parts of a project together, as discussed in Section 2.5. This fitting together can be done ahead of the research, or during the research, but either way it needs to be done. Just as Section 2.1 of this chapter stressed the pragmatic benefits of 'questions first – methods later' in maximising that fit, so this section stresses the

pragmatic benefits of beginning with research questions that are at least reasonably well developed.

To summarise: There is a continuum of possibilities, which is about prespecifying versus unfolding structure in the research. It applies to research questions, design and data. The issue is structure and its timing – when in the research is structure introduced? Prespecified research does it ahead of the empirical procedures. Unfolding research does it during them. As a general rule, at least a reasonable level of specificity in the initial research questions is desirable, though various factors need to be taken into account in particular situations. Chapter 4 will describe a model of research where considerable effort is invested in developing research questions ahead of the empirical work. But this is not the only model, and when research questions come later, they still require both the analytical development described in Chapters 4 and 5, and the matching with methods, design and data described in Section 2.5 of this chapter.

Chapter summary

Methodological theory is theory about methods, and involves philosophy. This is because methods are based on paradigms. A paradigm is a set of assumptions about the world.

The questioning of paradigms led to a prolonged quantitative-qualitative debate, characterised by either/or thinking. This was a very prominent feature of the paradigm wars among philosophers and methodologists, which took place in the 1960s, 70s and 80s.

Paradigm issues have more recently converged into positivism (mainly associated with quantitative methods) on the one hand, and interpretivism or constructivism (mainly associated with qualitative methods) on the other.

A research project can be paradigm-driven, where it begins with a paradigm, and develops research questions and methods from it, or pragmatic, where it begins with research questions which need answers, and chooses methods for answering them.

The purpose of substantive theory is to explain some substantive phenomenon of interest.

Description and explanation are two different levels of understanding of empirical data. Both are important, but the overall purpose of scientific research is explanation, not just description. This shows the importance of explanatory theory.

Theory verification research begins with a theory, develops hypotheses from this theory, and then tests the hypotheses against empirical data. By contrast, theory generation research starts with research questions and data, and aims to end with a theory which explains the data.

Good research has a close fit between the questions it asks and the methods it uses. A very good way to achieve this fit is for methods to follow from questions.

In prespecified research, the research questions and methods are preplanned, and the empirical part of the research implements these methods. In unfolding research, the questions and the methods are, to some extent at least, developed as the research proceeds. The difference is in the timing of the structure of the research. Which approach is 'better' needs to be determined in each particular research situation.

KEY TERMS

Methodological theory: theory about methods, and the philosophical assumptions which (necessarily) underlie any set of research methods

Paradigm: a set of assumptions about the world, and about what constitute proper topics and techniques for inquiring into that world. Paradigms have an ontological dimension (concerned with the nature of reality), an epistemological dimension (concerned with knowledge about that reality) and a methodological dimension (concerned with methods for building knowledge of the reality)

Positivism: the philosophical position that objective accounts of the world can be given, and that the function of science is to develop descriptions and explanations in the form of universal laws - that is, to develop nomothetic knowledge

Interpretivism: the philosophical position that people bring meanings to situations, and use these meanings to understand their world and influence their behaviour

Constructivism: the philosophical position that realities are local, specific and constructed, and are socially and experientially based, depending on the people holding them

Paradigm-driven research: research which begins with a paradigm, and develops research questions and methods from it

Pragmatic research: research which begins with research questions, and then chooses methods for answering them

Substantive theory: content-based theory, which aims to develop a set of internally consistent propositions to explain a substantive phenomenon of interest; substantive theory is explanatory

Description: using data to draw a picture of a situation, event, person (people) or something similar; focuses on what is the case

Explanation: accounting for a description, showing why and how events or situations have come to be what they are; focuses on why (or how) something is the case

Theory verification research: research which sets out to test a theory, by testing hypotheses derived from the theory; begins with theory

Theory generation research: research which starts with research questions and data, and aims to build a theory to explain the data; finishes with theory

Question-method fit: the need for internal consistency between the research questions asked, and the methods used for answering them; an important aspect of the validity of a piece of research

Pre-specified research: research which has a high degree of structure before the empirical work is done; research questions, methods and data are specified in advance

Unfolding research: research which does not have a high structure before empirical work begins; initial research questions may be loose and general, and more specific questions, methods and data are developed during empirical work

Exercises and study questions

- What is a paradigm? What are the three main dimensions of paradigms?
- What were the 'paradigm wars'?
- How are paradigms and methods connected?
- What is a paradigm-driven approach to research? What is a pragmatic approach to research?
- How do they differ?
- What would a *description* of the climate of (say) a London winter look like? What would an *explanation* of that climate look like? How are they different?
- For what sorts of topics and research questions would prestructured research be appropriate?
- For what sorts of topics and research questions would unfolding research be appropriate?

Further reading

- Anfara, V.A. and Mertz, N.T. (2006) *Theoretical Frameworks in Qualitative Research*. Thousand Oaks, CA: SAGE.
- Babbie, E. (2012) *The Practice of Social Research*. 13th edn. Belmont, CA: Wadsworth.
- Bailey, K.D. (2008) *Methods of Social Research*. 4th edn. New York: Simon & Schuster.
- Creswell, J.W. (2013) *Research Design: Qualitative and Quantitative Approaches*. 4th edn. Thousand Oaks, CA: SAGE.
- Lewins, F. (1992) *Social Science Methodology*. Melbourne: Macmillan.
- Little, D. (1991) *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder, CO: Westview.
- Marshall, C. and Rossman, G.B. (2010) *Designing Qualitative Research*. 5th edn. Thousand Oaks, CA: SAGE.
- Maxwell, J.A. (2012) *Qualitative Research Design: An Interactive Approach*. 3rd edn. Thousand Oaks, CA: SAGE.
- Miles, M.B., Huberman, A.M. and Saldaña, J. (2013) *Qualitative Data Analysis*. 3rd edn. Thousand Oaks, CA: SAGE.
- Neuman, W.L. (2010) *Social Research Methods: Qualitative and Quantitative Approaches*. 7th edn. Harlow: Pearson.
- Wolcott, H.F. (1994) *Transforming Qualitative Data: Description, Analysis, and Interpretation*. Thousand Oaks, CA: SAGE.

Notes

The 'paradigm wars' were especially vigorous in the field of education research. A good record of those 'wars', including the moves towards reconciliation and detente, can be found in a series of articles in *The Educational Researcher*, beginning in the 1970s.

The 'difficult' level is about precise definitions of the two terms, and about philosophical investigations into the concept of explanation – see, for example, Little (1991) and Lewins (1992).

Explanation itself is a complex philosophical concept. Another form of it is the 'missing links' form. Here, an event, or empirical generalisation, is explained by showing the links

that bring it about. Thus the relationship between social class and scholastic achievement might be explained by using cultural capital (Bourdieu, 1973) as the link between them. Or the relationship between social class and self-esteem might be explained by using the parent-child relationship as the link between them (Rosenberg, 1968: 54-82).

A nomothetic view sees generalised knowledge, universal laws and deductive explanations, based mainly on probabilities derived from large samples, and standing outside the constraints of everyday life. An ideographic view sees nomothetic knowledge as insensitive to local, case-based meanings, and directs attention rather to the specifics of particular cases. It prefers to see knowledge as local and situated (Denzin and Lincoln, 2011). The ideographic view thus points towards understanding and interpretation as important goals of research, alongside description and explanation.

Note also Atkinson's (1978) critique of that work, focusing on how suicide rates are constructed and what they mean.

CHAPTER 1

The Theory of International Politics

Therefore, the seeker after the truth is not one who studies the writings of the ancients and, following his natural disposition, puts his trust in them, but rather the one who suspects his faith in them and questions what he gathers from them, the one who submits to argument and demonstration, and not to the sayings of a human being whose nature is fraught with all kinds of imperfection and deficiency. Thus the duty of the man who investigates the writings of scientists, if learning the truth is his goal, is to make himself the enemy of all that he reads, and, applying his mind to the core and margins of its content, attack it from every side. He should also suspect himself as he performs his critical examination of it, so that he may avoid falling into either prejudice or leniency. (Ibn al-Haytham)¹

Since the end of World War II, debates about the relation between war and the state among political scientists in the United States have been dominated by a body of ideas commonly called Realism and the criticisms those ideas have provoked.² Nearly every author who wants to write something portentous about international politics either defends Realism, invents a new species of it, or uses it as a point of departure for some other “ism” that he or she wants to defend. Prominent among these alternatives to Realism have been Liberalism (including what has been called neoliberal institutionalism) and Constructivism.

Because mathematical models based on the theory of games have been used to evaluate the competing claims made by Realists and their critics, debates about Realism have become embroiled in controversies about game theory, the use of mathematics to describe human behavior, and something called “rational choice theory.” Some Constructivists have claimed that what is at stake is nothing less than fundamental issues in the philosophy of science or even something called “ontology.” The result has been not only to make the controversies provoked by Realism even more

1. Quoted in Sabra 2003, 54.

2. I will always capitalize the term *Realism* when I use it to refer to the academic doctrine that goes by that name among students of international politics.

difficult to settle but also to create confusion about what a theory of international politics might consist of or how to evaluate it.

Theories, Arguments, and Explanations

I will argue that the fundamental cause of the unproductive nature of these controversies has been the willingness of political scientists to tolerate incomplete arguments. Indifference to the validity of arguments is often justified by the claim that the issues raised in these debates are fundamentally empirical ones and that therefore they can only be settled by looking at the facts. If what I have to say is to be persuasive, I must show why this common view is mistaken. Let us begin, then, by looking at a few simple examples that will make this clear.

Arguments and Explanations

Suppose you went to the dog pound to look for an inexpensive dog and wanted to make sure that any dog you got had a friendly disposition, was good with children, and would not maul a passing stranger. Suppose the attendant assured you that a particular dog would have those qualities. Skeptical, you might ask, "How do you know that, and why should I believe it?" The attendant might reply that the dog in question was a Labrador retriever. "So?" you might reply, to which the attendant might respond that Labrador retrievers are friendly dogs and are good with children.

The attendant's answers to your questions can be interpreted as an argument, perhaps the simplest possible argument that actually conveys new information. It has two premises: "Labrador retrievers are friendly and good with children" and "This dog is a Labrador retriever," from which it follows that "This dog will be friendly and good with children," which is what you wanted to know. The conclusion "follows from" the premises only because if one accepts the premises and denies the conclusion one would have contradicted oneself, which is why if one believes the premises one must also believe the conclusion. Arguments that have this property are called valid arguments, and reasoning from premises to conclusion in this way is commonly called "deductive reasoning."

However, this little argument would satisfy you only if you were confident that both of the premises were true. If one or both were not true, the argument would remain valid but the conclusion might be false. Suppose, then, that you asked why you should believe that this dog was a Labrador retriever—this is, after all, the dog pound. The attendant might reply that Labrador retrievers had certain recognizable characteristics

such as a large, square head, short hair, a wide chest, and a friendly disposition, and this dog had those characteristics.

At a glance this looks like a deductive argument just like the first one: the first premise is that Labrador retrievers have certain recognizable characteristics, and the second is that this dog has all those characteristics. But if so, the argument is not valid, because it does not follow from these two premises that the dog is a Labrador retriever. Such an argument would be an example of a logical fallacy called “affirming the consequent” and therefore could not provide the assurance you were looking for.

But this would be a misunderstanding of the attendant’s reasoning. The attendant is saying that the hypothesis that the dog is a Labrador retriever would explain its appearance, and thus its appearance gives us reason to believe that it is a Labrador retriever. There is a deductive argument here, but its premises are that “All Labrador retrievers have certain recognizable characteristics” and “This dog is a Labrador retriever,” from which it would follow, if true, that this dog would have the properties of a Labrador retriever. But this is something that one does not have to be persuaded of, since the dog can be inspected directly. The question is, rather, what sort of dog is it? And the reasoning is that, since these premises, if true, would imply that the dog would have the appearance that it does have, the fact that it has that appearance is evidence that the premises are true. This is an example of what is commonly called “inductive reasoning,” and the problem of induction is to figure out what justifies an inference of this sort.³

But we do not require a justification for reasoning in this way to do it.⁴ What is important here is, rather, the fact that the inductive inference from the dog’s visible characteristics to its breed is made possible by a deductive inference from the breed to a dog’s visible characteristics: if the breed could not explain its appearance, then the breed could not be inferred from the appearance.⁵ The problem is that there are other possible explanations of the dog’s appearance, some of which might imply that it would be dangerous, and that is why inductive inference requires not just identifying a possible explanation of the facts but also supplying reasons to believe that

3. Probability theory provides a plausible answer to that question, since such an inference can be shown to be an application of Bayes’s rule. For a recent discussion by a philosopher, see Howson 2000. For a discussion by a physicist, see Jaynes 2003.

4. A person who did not engage in inductive reasoning would not soon survive, since unlike most animals the behavior encoded in the genes of human beings is inadequate for humans to cope with their environment. Thus there is reason to believe that human beings are endowed with a propensity to engage in it—even infants do it (Ruse 1998; Gopnik, Meltzoff, and Kuhl 1999).

5. In the Bayesian interpretation of inductive inference, the deductive argument tells us that the conditional probability of the conclusion being true, given the truth of the premises, is one.

that explanation is better than other possible ones. Thus inductive inference is sometimes said to be “inference to the best explanation.”⁶

Similarly, if you asked why you should believe that Labrador retrievers are good with children, you might be told that many people had had such dogs as pets and this was their uniform experience. Since the hypothesis that all Labrador retrievers are good with children would explain the fact that everyone who had had them as pets found them to be good with children, that fact is evidence that the hypothesis is true. However, a cautious person might wonder if there were other possible explanations of this fact.⁷

In spite of the fact that they are almost trivially simple, these examples illustrate how claims to knowledge are justified. More complex examples could be found in detective stories, murder trials, investigations of the causes of plane crashes, troubleshooting procedures for automobile mechanics or people who service computers, and throughout the natural sciences.

These examples also illustrate the fact that whether we are reasoning from premises to conclusions, or from observable facts to possible explanations of those facts, what is commonly called logical validity is necessary if our reasoning is to affect our beliefs: if the confidence we place in some premises is to be transferred to a conclusion then the conclusion must be implied by the premises, and if some explanation is to be supported by the facts then the facts must be implied by the explanation.⁸ In these examples satisfying this requirement is so easy that it is possible to overlook it. Unfortunately, in even slightly more complicated situations it is possible to think one has satisfied it when one has not. We will see that this is true of much that has been written about international politics.

Science, Causes, Variables, and Theories

People who are interested in making the study of politics scientific often consult works on the philosophy of science to tell them what a science is supposed to consist of.⁹ Many come away with the idea that the aim of sci-

6. C. S. Peirce called inference to the best explanation “abduction,” and many writers restrict the word *induction* to inference from a sample to a population. My usage conforms to recent literature that argues that the logic of inference in the two cases is the same (Harman 1965; Thagard 1978; and Lipton 1991).

7. See, for example, Malcolm Gladwell’s (2006) comparison between generalizations about the behavior of various breeds of dogs and the development of profiles of potential suspects by police.

8. Or at least the explanation must imply that the facts should have been expected with some probability.

9. For an engaging discussion of this sort of thing, see George Homans’s (1984) autobiography.

ence is to identify causal regularities. Since a curve defined by an equation in which a dependent variable is a function of one or more independent variables seems to be a way of representing a causal regularity, statistical techniques for fitting such a curve to numerical data are often the standard by which the study of politics is judged. As a result, even people who do not use statistics couch their explanations in terms of independent and dependent variables, and attempts to explain individual events are commonly described as “small *n* studies” or “case studies” or are said to commit the statistical sin of “selection on the dependent variable.”¹⁰

But the philosophy of science is mainly about the problem of induction, which is a problem for philosophers, not for scientists, and it is a serious mistake to think that one might find in it a blueprint for doing science. Moreover, the word *science* is not well defined, and the only thing that all the fields of study commonly called sciences seem to have in common is that (1) they all reward people for showing that existing explanations of the phenomena described by the field fail to meet the standards for justifying claims to knowledge discussed above, (2) they give even greater rewards to people who construct nonobvious explanations that survive attempts to discredit them in this way, and (3) they require scholars to make their work as easy to criticize as possible by making the reasoning that supports it transparent (Ziman 2000). Thus a plausible definition of *science* is just that it refers to any enterprise in which scholars compete with each other in constructing nonobvious explanations of the phenomena they study that can withstand concerted attempts to discredit them.¹¹

It is not really clear what a “causal regularity” is, but by any ordinary definition of that term much of what is commonly called “science” is more concerned with explaining regularities than with identifying them. Science does not tell us that the sun rises in the east and sets in the west, that the sky is blue, or that the days are longer in the summer than in the winter—it explains those facts. Or, to take a less obvious example, cooks discovered that whipping egg whites into a meringue works best in a copper bowl. The physical sciences explain why that is true. The causal regularity was discovered by cooks; “science” explains it.¹²

10. For an influential example, see King, Keohane, and Verba 1994.

11. Many people believe that the philosopher Karl Popper defined at least one test for distinguishing between science and nonscience, and that is the requirement that propositions be falsifiable. However, while genetics may one day provide a means of falsifying the proposition that an anonymous dog at the dog pound has the genes that give Labrador retrievers their characteristic disposition, that is still, apparently, not possible. But that does not make inferences from its appearance to its breed meaningless or unjustifiable (Howson and Urbach 1993; Howson 2000; Ziman 2000, 226).

12. See Derry 1999, 4–6. Note that the explanation also increases our confidence that what cooks say is not just a superstition. See also the physical explanations in Chandrasekhar 1998.

However, what is called science does not just explain regularities, whether causal or otherwise; it also explains unique events, for example, where the HIV virus came from and, if it came from chimpanzees, how it got from animals to people. It is absurd to think that this is an example of a “small n study” that would be assisted by an increase in the size of the sample. Rather, the problem is to identify possible explanations of what happened and then to see how many of the known facts each explains.¹³

Consider the problem of explaining plane crashes. People charged with that grisly and important task often know very little about what happened: all the eyewitnesses may be dead, and the plane itself may be smashed to pieces and not fully recoverable. In addition to what can be recovered from the wreckage, they know the pattern formed by the debris, some of the weather conditions when the plane crashed, perhaps some or all of the data on the flight recorders, and sometimes a recording of radar images of the trajectory of the plane as it crashed. Their problem is to find an explanation that accounts not only for the plane crash but also for more of the other information at their disposal than does any other explanation. The plane crash is not a dependent variable whose variation might be accounted for by one or more independent variables. It is a fact, and what is wanted are some propositions from which, if true, that fact could be derived. Thus one must reason backward from what is known to what is unknown, and the only evidence there is for the truth of the investigators’ conclusion is that it explains the known facts. However, since more than one explanation might account for those facts, an effort must be made to avoid settling on the first one that comes to mind.

One broad category of explanations for plane crashes falls under the heading of “pilot error.” One might wonder if there is a relationship between such things as pilot training or flight schedules and pilot errors serious enough to cause plane crashes. If so, pilot error might be taken as a dependent variable, and one might test for a relationship between it and such independent variables as training procedures or frequency of flying. But such a relationship, if found, could not be said to explain the plane crashes. And to measure the dependent variable one must first have explained all the plane crashes individually, in order to know which ones were the result of pilot error and which were the result of mechanical failure or some other cause—the fact that all plane crashes are in some sense the same does not mean that they can all be explained in the same way. Moreover, to explain any plane crash one must first be able to explain why planes are able to fly.

Contrast this discussion of flight failure with the recent political sci-

13. See the account in Kolata 2001.

ence literature on “state failure.”¹⁴ In 1994, at the request of Vice President Al Gore, the U.S. government established a State Failure Task Force, composed of prominent social scientists, whose purpose was to determine what conditions were associated with the failure of states. It found three independent variables that accounted for most of the state failures—infant mortality, trade openness, and democracy—and on that basis made a number of policy recommendations to the U.S. government. What is one to make of such a study?

One criticism one might make is that it has not properly identified the dependent variable or that some of the independent variables are really part of the dependent variable (King and Zeng 2001, 654–55). But more serious problems are implied by our plane crash example. It is not clear how one could explain state failure without being able to explain state success, which we are far from being able to do. Moreover, as with plane crashes, there is no reason to think that all state failures (whatever that might mean) can be explained in the same way. And finally, in explaining the “failure” of any state, the problem is not to find independent variables that would account for the variation in some dependent variable but to find a set of propositions from which the facts of interest could be derived. One of those facts would be the relation between the independent and dependent variables reported by the State Failure Task Force.¹⁵

Some people would say that this implies that to understand state failure we need a theory of state success. But the word *theory* means so many things and has been used in so many different ways by political scientists that that would not be very informative. Any conjecture can be called a theory, and it would not be at all surprising to find a political scientist referring to studies of state failure of the sort just summarized as “state failure theory.” Everyone aspires to “do theory,” and it is often said that there are many different ways of doing it and we should be tolerant of all of them.¹⁶

However, while there may be many ways of “doing theory,” there are

14. See especially the discussion in King and Zeng 2001.

15. Confusion between the relation between independent and dependent variables in a regression equation, on the one hand, and the relation between premises and conclusion in an explanatory argument, on the other, is common in political science, and many political scientists claim that a theory just consists of a specification of the relation among a set of variables. See, for example, Van Evera 1997, 7–48.

16. This usage reflects the influence of postmodernist writings on the study of literature. See the account in Culler 1997, chap. 1, titled “What Is Theory?” Culler writes: “In literary and cultural studies these days there is a lot of talk about theory—not theory of literature, mind you; just plain ‘theory’. . . . ‘Theory of what?’ you want to ask. It is surprisingly hard to say. . . . Sometimes theory seems less an account of anything than an activity—something you do or don’t do.”

not many ways of constructing valid arguments that can serve as explanations of observed facts. What is wanted is not just anything that might be called a theory but an explanation from which the facts in question can actually be derived. It is that sort of theory that is the subject of this book.

Causality and Meaning

Many people would say that human behavior is too unpredictable for such explanations to be possible. However, after saying that, such people will often literally bet their lives that what they have said is not true by driving a car at seventy miles an hour down a highway while separated from cars traveling at the same speed in the opposite direction only by a painted yellow line. And in buying the car they drive they will have bet a lot of money that wherever they go there will be people willing to supply them with oil and gasoline to keep it running and to fix it when it breaks down. Human behavior is, in fact, very predictable, and if it were not, social organization would be impossible.

Ants seem remarkable to us because their social organization resembles that of humans, and they engage in complex forms of cooperation that look very much like war, gardening, and the domestication of animals.¹⁷ In their famous book about ants, Hölldobler and Wilson say:

The study of ant social organization is by necessity both a reductionistic and a holistic enterprise. The behavior of the colony as a whole can be understood only if the programs and positional effects of the individual members are both specified and explained more deeply at the physiological level. But such accounts are still far from complete. The information makes full sense only when the colonial pattern of each species is examined as an idiosyncratic adaptation to the natural environment in which the species lives. (1990, 3)

If one substitutes “psychological” for “physiological” in this quotation, one gets something very close to the following statement by the German sociologist Max Weber:

Interpretive sociology considers the individual . . . and his action as the basic unit, as its “atom.” . . . In general, for sociology, such concepts as “state,” “association,” “feudalism,” and the like, des-

17. Comparisons between the social organization of humans and the social organization of insects can be found in both Hobbes and Aristotle. For a recent development of the theme, see Skyrms 2004, xi–xiv.

ignite certain categories of human interaction. Hence it is the task of sociology to reduce these concepts to “understandable” action, that is, without exception, to the actions of participating individual men. (Gerth and Mills 1946, 55)

These two quotations touch on two issues that are the source of vast amounts of unnecessary conflict and confusion. These issues have been revived by Constructivists in their quarrel with Realism.

One issue is whether the social sciences should be “reductionistic” or “holistic,” to use the terminology of Hölldobler and Wilson. What they say of this in connection with the study of the social behavior of ants seems obviously true of human beings as well: it must be both.¹⁸

The other is whether substituting “psychological” for “physiological,” in the quotation from Hölldobler and Wilson, implies that the scientific study of society is impossible. The basis for claiming that it does is that physiology is about what causes what, while psychology is, as Weber said, about meaning.¹⁹ However, as already mentioned, it is not really clear what “causality” means, and if what we are interested in is finding non-obvious explanations for what happens, then it makes perfectly good sense to speak of social science—though, contrary to what many believe, doing so tells us little about how to proceed. And once it becomes clear that we are interested not simply in whether some dependent variable can be made to jiggle by yanking on some independent variable but in why that might be true, one can see that the same criteria for evaluating explanations apply to both realms.

We are curious about how to explain the complex social behavior of ants because it seems so much like what human beings do. If we are to find nonobvious explanations of human social behavior, we must learn to become as puzzled by what humans are capable of doing as we are by the ants. Thus, instead of looking for new, unsuspected regularities that might be found in human behavior, it might be useful to begin by thinking about whether we can explain the regularities in it that are as familiar to us as the rising and setting of the sun or the progress of the seasons and that we all take completely for granted. For example, instead of being puzzled by what is now called state failure, we should be puzzled by state success, which is actually the rarer phenomenon if, by the word *state*, we mean the

18. See the discussion of controversies about this issue among biologists by Edward O. Wilson (1994). For a very interesting discussion of this issue in the social sciences by an evolutionary biologist, see David Sloan Wilson (2002).

19. For an extended argument of this sort, see Winch 1958, which claims that “the conceptions according to which we normally think of social events are logically incompatible with the concepts belonging to scientific explanation” (95).

modern states, whose leaders are now so concerned about state failures around the world.²⁰

Nonobvious explanations are, like nonobvious theorems, not obvious! There is nowhere to look up what they might be and no one to tell us in advance what will work and what will not. Unlike natural scientists, social scientists have the advantage of being able to think like the people whose behavior they want to explain.²¹ Moreover, it is counterintuitive to think that one could be part of a social organization without already understanding it. But one can know the important elements of a good explanation without seeing their implications, especially if they imply an explanation that is different from one that everyone already accepts. In Darwin's time, the ideas of Thomas Malthus were widely known and what animal breeders did was familiar to nearly everyone. However, only two people saw that those ideas together implied that complex organisms could have developed without an intelligent designer: A. R. Wallace and Charles Darwin. They would not have done so had they stuck with what seemed obviously true to everyone else. It is also unlikely that they would have done so had they first consulted a treatise on the philosophy of science, or a statistics textbook, before proceeding.²²

Models: Method or Madness?

A model is just something that is used to represent something else, like a model airplane. Everyone who has used a map or a house plan or an architect's drawing has used a model. The purpose of such models is to facilitate inferences about the thing that is modeled that would otherwise be difficult. You could try to figure out how to landscape your yard or arrange the furniture in your new house just by standing in the middle of it and thinking about how it will look, but you might find it easier to work with a drawing. Similarly, you could give your guests complicated verbal instructions about how to find your house, but it might be more effective to give them a map and let them draw the proper inferences from it.

Whenever we use models such as these we have to worry whether con-

20. See the discussion by Paul Seabright (2004) of how remarkable and puzzling the development of large-scale political and economic organizations by human beings really is.

21. This is not a trivial point, and it is important not just for social science but for human social organization as well (Ziman 2000, 107–9). See the fascinating discussion of the psychological literature on this subject in Baron-Cohen 1995.

22. Darwin's reasoning was criticized by some of the leading philosophers of his time for failing to satisfy appropriate standards of inductive inference—see the discussion in Hull 1973. For an account of the development of Darwin's ideas, see Mayr 1991, 68–89. See also Press and Tanur 2001, which contains accounts not only of the development of Darwin's ideas but of many other important scientific ideas as well. For further discussion of the problem of explanation in the social sciences, see the witty and enjoyable analysis in Homans 1967.

clusions that we reach that are true of the model also apply to the thing or things that the model represents. If a drawing of one's house or lawn is not drawn exactly to scale, then things that \nless in the drawing won't \nless in "the real world," and if roads that look straight on a map are really very crooked, then it may take longer to get to your house than your friends thought. There are always differences of this sort between models and the things they represent, and the question therefore is not whether the model is completely accurate (no model is or can be, or it would not be a model) but whether it is accurate enough for the purpose at hand. A map that is good enough to enable people to \nless your house might not be good enough to determine how much fiber-optic cable to buy if a company plans to wire your neighborhood or to plot the path of a cruise missile.²³

The same is true of models of nonphysical things. Formal or symbolic logic, for example, is a system of arbitrary symbols and rules for manipulating them that was designed to represent logical inference. Since the rules for manipulating the symbols are absolutely clear, it is often easier to prove theorems by using them than it is by using words. However, that can lead to controversies about whether theorems that are true in this symbolic language always carry over to the ordinary language that everyone actually thinks in (Strawson 1952). And what is nowadays called rational choice theory, in disputes about theories and methods among political scientists, is really just a way of constructing mathematical models of people's choices, which can lead to similar controversies (Wagner 2001).

Since reasoning about models instead of the real thing can be misleading, there has to be a good reason for doing it. And since we explain people's choices all the time without constructing models of them, the whole idea may seem ridiculous. There are three main ways in which explanations involving human choices can become complex enough that models of them can be useful. One is that the consequences of the choices of many people taken together may not be obvious and may then interact with people's subsequent choices. This is what happens in markets and in electoral systems with competing political parties. A second is that individuals may be faced with uncertainty about the consequences of their choices, so their choices are not implied in any straightforward way by their preferences over \nless nal outcomes. And a third is that individuals' choices may be interdependent, in that what one person will choose depends on his or her expectations about how one or more other people will choose and vice versa. It is not possible to understand international

23. For a discussion of such issues concerning models in the natural sciences, with examples, see Derry 1999, 69–88. An excellent introduction to the use of models in the social sciences can be found in Lave and March 1993. For an introduction to the use of mathematical models in studying international politics, see Powell 1999.

politics without confronting all these problems, which is why mathematical models have become so important in the study of it.

Formal models have helped us think much more clearly about many of the questions debated by students of international politics. However, what one gets out of a formal model depends on what one puts into it, and therefore game theory is not a ready-made theory of international politics, and no formal model can compensate for a poorly framed question. Many criticisms of formal models wrongly attribute the problems they identify to the use of mathematics, when they are instead the result of the way the problem has been formulated. The purpose of this chapter and the next one is to look carefully at how the questions debated by students of international politics came to be formulated in the way that they have been. I will then argue that they need to be reformulated.²⁴

A Guide for the Reader

One impediment to settling the issues raised by Realists and their critics is that it is not entirely clear what Realism is. There is now an embarrassment of Realisms. There is classical Realism, neoclassical Realism, structural Realism (aka Neorealism), human nature Realism, defensive Realism, and offensive Realism, and it may be undergoing further mutations as I write. Thus it will be necessary to figure out not only what is distinctive about each of the main varieties of Realism but also what, if anything, they all have in common.

Because I am interested in evaluating the current state of the field and not simply summarizing its historical development, I will begin with recent varieties of Realism and work backward from there, concluding with an examination of the origins of Realism. Along the way I will discuss the criticisms of Realism that have been made by its main competitors. Following the guidelines laid down earlier, I will not try to summarize everything that writers who identify themselves with these “isms” have said but will instead try to identify the main theses they have advanced and the arguments they have offered in support of them. In the next chapter I will examine the origins of these ideas in early modern European political thought, a subject that is frequently discussed by contemporary defenders of these competing doctrines but that they have often misinterpreted.

The purpose of these two chapters is to show that all these “isms” are

24. Unfortunately, defenders of formal models have contributed to confusion about their significance by saying that their purpose is to help us think consistently. To be inconsistent is to contradict oneself. The problem with the arguments made by students of international politics is not that they have been self-contradictory but that they have been invalid, that is, believing their premises and denying their conclusions would *not* be inconsistent.

collections of bad answers to important questions. An understanding of what is wrong with these answers will, I hope, lead to a better understanding of the questions, which are much broader than is commonly assumed by modern-day Realists and their critics. Beginning in chapter 3 I will investigate what contemporary scholarship has to say about these questions.

Every chapter prior to chapter 6 will lead to a new set of questions, which will be the focus of the following chapter. Following chapter 6 I will try to summarize the implications of the preceding chapters.

Offensive Realism

A recent version of Realism that has received a lot of attention is offensive Realism, and its main proponent is John Mearsheimer.²⁵ The main thesis of offensive Realism is that even states that want only to be secure act aggressively, because the international system forces them to do so.

This situation, which no one consciously designed or intended, is genuinely tragic. Great powers that have no reason to fight each other—that are merely concerned with their own survival—nevertheless have little choice but to pursue power and to seek to dominate the other states in the system. (Mearsheimer 2001, 3)

Thus whatever the nature of the component states, international politics “has always been a ruthless and dangerous business, and it is likely to remain that way” (2).

Unlike many writers, Mearsheimer actually lists five properties of international politics that together, he claims, logically imply this conclusion (30–32). They can easily be stated in the form of five premises:

Premise 1: *There is no world government.*

Premise 2: *All states are capable of using force against other states.*

Premise 3: *No state can ever be certain that another state will not use force against it.*

Premise 4: *All states seek to maintain their territorial integrity and domestic autonomy.*

Premise 5: *States are rational actors.*

After listing these assumptions, Mearsheimer says:

25. The distinction between offensive and defensive Realism is apparently due to Jack Snyder (1991, 10–13).

none of these assumptions alone dictates that great powers as a general rule *should* behave aggressively toward each other. There is surely the possibility that some state might have hostile intentions, but the only assumption dealing with a specific motive that is common to all states says that their principal objective is to survive, which by itself is a rather harmless goal. Nevertheless, when the <ve assumptions are married together, they create powerful incentives for great powers to think and act offensively with regard to each other. (32)

Thus Mearsheimer claims to have derived a strong, nonobvious conclusion from premises whose truth it would be hard to deny.

However, he makes no attempt to show that his conclusion follows from these premises. Had he done so, it would have been more obvious that, while his premises are clearly stated, it is far from clear what the conclusion actually is. In the passage quoted earlier, he says that, because of these properties of international politics, states “have little choice but to pursue power and to seek to dominate the other states in the system.” And in the passage just quoted he says that these properties “create powerful incentives for great powers to think and act offensively with regard to each other.” But, while it is clear that he thinks that these statements may be true even if the only objective of states is to survive, it is not really clear what they mean.

One possibility that is consistent with what he says is the following statement:

Conclusion 1 (Mearsheimer?) *Two states may go to war with each other even though they both want only to survive.*

But, while we may be able to think of circumstances in which this statement would be true, it does not follow from these premises.

Another possible interpretation of what he says is something like this:

Conclusion 2 (Mearsheimer?) *Even a state that wants only to survive will want to have more powerful military forces than all other states combined.*

But not even this statement follows from his premises.²⁶

Because Mearsheimer never makes clear exactly what he thinks follows from his premises, the fact that his argument is not valid is not as

26. Note that the issue is not whether international politics has in fact often been a “ruthless and dangerous business,” as Mearsheimer says, or even whether these hypothetical conclusions are sometimes true, but what would explain their truth and therefore whether they must be true as long as Mearsheimer’s premises are true.

obvious to the reader as it might be. Impressed by the seemingly obvious truth of the premises, readers may be unjustifiably impressed by the argument itself.²⁷

This is the key to explaining the dominant role that Realism has played in the study of international politics: it claims to derive strong conclusions about the behavior of states from properties of international politics that are difficult to deny. But the claim is unjustified, not just in Mearsheimer's case but in others as well.²⁸

Defensive Realism

No Realist of any type would quarrel with any of Mearsheimer's core premises. Nonetheless, as Mearsheimer points out, other Realists do not accept his conclusion (Mearsheimer 2001, 19–22). (This would be more clearly true if it were clearer what his conclusion actually is.) But if Mearsheimer's claim about the implication of these premises is correct, they must have made a mistake in their reasoning. Yet nowhere does he say exactly what this mistake is.

For example, he says of Kenneth Waltz, the most prominent of the scholars he identifies as a “defensive realist,” that there is a “status quo bias” in his theory, leaving one with the impression that Waltz merely assumes that states only want to protect the status quo, without seeing that his own assumptions imply something different (20). However, Waltz did not merely assume that states were not interested in aggressively expanding; he argued that even if they were inclined to expand, the very features of the international system that Mearsheimer describes would lead them not to do so. Mearsheimer simply ignores Waltz's argument, asserts a different conclusion, and gives it a distinctive name to emphasize the nature of the difference.²⁹

27. On the cover of the paperback edition of Mearsheimer's book, Samuel Huntington is quoted as saying of it: “All serious students of international affairs will have to come to grips with its argument.” On the back, the *Economist* is quoted as saying that Mearsheimer “demolishes all the main components” of the “happy vision” of international politics that emerged at the end of the cold war.

28. Mearsheimer's third premise is actually far more questionable than it appears to be. I will examine it in chapter 6. But first we must determine why it is important.

29. Schweller (1996) also claims that Waltz assumed that states were only interested in security, not expansion. Waltz is frequently unclear, but he often and explicitly says otherwise in more than one place. “Beyond the survival motive,” he wrote, “the aims of states may be endlessly varied; they may range from the ambition to conquer the world to the desire merely to be left alone.” What he assumed was just that “[s]urvival is a prerequisite to achieving any goals that states may have” (Waltz 1979, 91). Otherwise the argument summarized subsequently would have been unnecessary, and his reasoning would have been not merely invalid but absurd.

Waltz's argument, while no more valid than Mearsheimer's, makes even more apparent the fact that Mearsheimer's argument is not valid either. It rests on the fact that, when there are more than two states in the world, one state cannot expand without giving another an opportunity to expand as well, in which case an increase in one state's absolute level of military capabilities might leave it relatively worse off than it was before. Hence, Waltz claims, the nature of international politics forces all states to focus on preserving their own independence by forming balancing coalitions rather than on maximizing their own power. Thus even expansionist states

cannot let power, a possibly useful means, become the end they pursue. The goal the system encourages them to seek is security. Increased power may or may not serve that end. (Waltz 1979, 126)

Mearsheimer claims, however, that the nature of international politics forces even satisfied states to be aggressive in order to maximize their power. Yet both claim to have derived their conclusions from exactly the same properties of international politics.

Mearsheimer, as we have seen, makes no attempt to show that his conclusion follows from these properties. This is Waltz's attempt to justify his conclusion:

Because power is a means and not an end, states prefer to join the weaker of two coalitions. . . . If states wished to maximize power, they would join the stronger side, and we would see not balances forming but a world hegemony forged. This does not happen because balancing, not bandwagoning, is the behavior induced by the system. The first concern of states is not to maximize power but to maintain their positions in the system.

Secondary states, if they are free to choose, flock to the weaker side, for it is the stronger side that threatens them. (Waltz 1979, 126–27)

This is the entirety of Waltz's argument in support of his claim that anarchy leads to "the recurrent formation of balances of power" (119). But it is not valid. A state that joins with the more powerful of two states to fight the third will confront a more powerful adversary after victory, but if it allies with the weaker state instead it will be less likely to be victorious. Without more information we cannot say what it should be expected to do.³⁰

30. For an extended analysis of the question raised by this passage from Waltz's book, see Powell 1999, chap. 5.

In discussing the difference between offensive and defensive Realism, Snyder says:

anarchy is not in itself sufficient to predict an expansionist security strategy. Realist scholars argue that the normal response to threat is to form a balancing alliance. Therefore states should expect that expansion will reduce their security insofar as it threatens other states and provokes an opposing coalition. (1991, 22)

This may be what some people who have identified themselves as Realists say. But there is no more reason to believe that it is true than there is to believe Mearsheimer's claim for the opposite view.

Structural Realism

The claim that propositions about the behavior of states can be deduced from properties of the state system is the most basic idea in what is often called structural Realism, or Neorealism. This claim was advanced originally by Kenneth Waltz, but it is accepted by Mearsheimer as well, and therefore offensive and defensive Realisms are both varieties of structural Realism. What is not commonly recognized is that the mere fact that structural Realists disagree about which of these views is correct is enough to call structural Realism itself into question.³¹ However, there is more to structural Realism than just the question of whether the international system makes states expansionist or instead curtails any expansionist tendencies they might have.

There are two important structural attributes of a state system, according to Waltz. All state systems, he claimed, are alike in having anarchic rather than hierarchical structures. However, there are also structural differences among anarchic systems as a result of differences in the distribution of power among the constituent states (Waltz 1979, 79–101). The disagreement between defensive and offensive Realists is one of the questions that arise about the properties of all systems with an anarchic structure. The other main issue raised by structural Realism is what differences

31. Instead, people who are reluctant to conclude that published works might just be wrong often try to save structural Realism by claiming that disagreements among Realists are the result of differing tacit assumptions made by authors who disagree. However, not only is there no textual basis for doing this, but typically no attempt is made to show that the respective arguments might be valid even with the extra assumptions. For an example, see Brooks 1997.

among anarchic systems can be attributed to differences in the distribution of power within them.³²

Prior to Waltz, most writers about international politics focused on the question of whether power among states was distributed equally or unequally.³³ Since Waltz's writing, nearly everyone has focused on whether systems were characterized by a bipolar or a multipolar distribution of power. There have been controversies about what sort of state behavior can be expected in each type of system and also about whether other types of "polarity" are possible and what effect they might have. I will have more to say about the effects of anarchy later. First let us consider the polarity of state systems.

Waltz's distinction between bipolarity and multipolarity was a product of debates about how to understand the cold war. During the period between 1945 and 1950, when it gradually became clear that World War II would not, as most people had expected, end like World War I with a comprehensive peace settlement but would lead instead to a protracted conflict between the United States and the Soviet Union, there were two features of the international situation that many people found deeply disturbing. One was the development of nuclear weapons, whose existence was unknown to most people until 1945. The other was the fact that after World War II the major powers seemed to be organizing themselves into two hostile coalitions separated by unbridgeable ideological differences. At some point this second feature of postwar international politics began to be referred to as "bipolarity."

The combination of nuclear weapons with bipolarity led many people to fear that civilization itself was threatened. It is ironic that by the end of the cold war many people had concluded that it was precisely those two features of international politics that had turned the cold war into what some people now call "the long peace" (Gaddis 1987). While Waltz originally accepted the view that nuclear weapons were very dangerous, he was almost single-handedly responsible for convincing many people that bipolarity was good, not bad—good enough, in fact, to compensate for the dangers posed by nuclear weapons (Waltz 1964).³⁴

Waltz claimed that people who were concerned about the polarization of the world into two nuclear-armed camps as a result of the cold war had misunderstood the situation. What had happened was not that the world had divided into two cohesive alliances of the traditional sort but rather

32. While Waltz was responsible for the idea that international systems could be characterized by their structures, it is important to recognize that a very large part of what Waltz said about the effects of a system's structure on the behavior of states within it was first said by John Herz. Herz's discussion is often clearer than Waltz's. See especially Herz 1959.

33. See, for example, Claude 1962, chaps. 1–3; Sheehan 1996.

34. Later Waltz (1981) also argued that nuclear weapons were actually a good thing too.

that the number of great powers had been reduced to two. And a world in which there are just two great powers, he claimed, is less prone to war than a world in which there are more than two (Waltz 1964; 1979, 168–70). He called a world with just two great powers a bipolar world and a world with more than two great powers a multipolar world, and the names stuck.³⁵

Why did Waltz think that a bipolar world was less war-prone than a multipolar world? Because, he claimed,

States are less likely to misjudge their relative strengths than they are to misjudge the strength and reliability of opposing coalitions. Rather than making states properly cautious and forwarding the chances of peace, uncertainty and miscalculation cause wars. . . . In a bipolar world uncertainty lessens and calculations are easier to make. (Waltz 1979, 168)

Because alliances are important in a multipolar world but are not in a bipolar world, in other words, the sorts of miscalculations that lead to war are less likely in a bipolar world.

Is this a valid argument? Let us try to reconstruct it. Clearly one of Waltz's premises is just a definition:

Premise 1 (Waltz) Definition: *A bipolar world is one in which there are two great powers, and a multipolar world is one in which there are more than two great powers.*

Another premise is:

Premise 2 (Waltz) *Miscalculations of the relative strength or behavior of opposing states or coalitions of states can cause wars to occur.*

A third is:

Premise 3 (Waltz) *States are less likely to miscalculate the strength or behavior of states than of opposing coalitions.*

And the conclusion is:

Conclusion 1 (Waltz) *War is less likely in a bipolar world than in a multipolar world.*

35. However, the distinction, as well as the terminology and much of the argument Waltz gave for its significance, had earlier been introduced into the literature by John Herz (1959).

It should be obvious that this is not a valid argument.

What is missing? At the very least we need the following two additional premises:

Premise 4 *The only miscalculations that can lead to war are miscalculations about the relative strength or behavior of the great powers.*

Premise 5 *There is no other possible cause of war that might be more likely to occur in a bipolar world than in a multipolar one.*

If we believe the fourth premise, then reducing the number of great powers to two implies that incorrect expectations about the behavior or performance of coalitions cannot lead to war. And if we believe the fifth then there is no other possible factor that might influence the likelihood of war that we need to be concerned about. However, it is far from clear why we should believe either premise, and Waltz does not say why we should.

Waltz's silence on the fifth premise is an indirect consequence of the more general fact that he has virtually nothing to say about why war occurs at all (his second premise is justified by a reference to an earlier edition of Blainey 1988). This is a point that I will return to later. A close reading of what he says about bipolarity, however, will show that his silence on the question raised by the fourth premise is apparently the result of confusing a reduction in the number of *great powers* in the world with a reduction in the number of *states*. Clearly when there are only two states in the world, uncertainty about who will ally with whom in a conflict between them cannot arise. As Waltz put it:

Systems of two have qualities distinct from systems of three or more. What is the defining difference? . . . Where two powers contend, imbalances can be righted only by their internal efforts. With more than two, shifts in alignment provide an additional means of adjustment. (1979, 163)

Waltz claimed, in effect, that a reduction in the number of great powers to two was equivalent to a reduction in the number of states in the world to two. But nowhere does he say why this should be true.

Indeed, he could not possibly say that, because he never defines what a great power is. He expresses impatience at the question, saying that "one finds general agreement about who the great powers of a period are, with occasional doubt about marginal cases," but admits that "[w]e should not be surprised if wrong answers are sometimes arrived at." "The question," he says, "is an empirical one, and common sense can answer it" (Waltz 1979, 131). In fact, however, common sense cannot answer the question,

since the term *great power* has no standard meaning. Definitions without arguments are often pedantic, but arguments without definitions are often not valid. That is why mathematicians are so picky about definitions.³⁶

Thus the seeming plausibility of Waltz's reasoning about the difference between a bipolar and a multipolar world was the result of his equivocation between the number of states in the world and the distribution of power between or among them. As a result, he showed neither that during the cold war there were only two great powers nor, if that were true, that that fact would have the consequences he claimed for it. Moreover, his claim that bipolarity rendered other states unimportant to the United States and the USSR during the cold war would, if true, make it impossible to understand why the cold war occurred at all.³⁷

It is a remarkable fact that, in spite of all the discussion and debate about bipolarity and multipolarity, not to mention the possible consequences of "unipolarity" since the end of the cold war, neither Waltz nor anyone else has ever specified what the "polarity" of an international system refers to. And therefore no one has ever presented a valid argument in support of the claim that states behave differently in systems with different polarities.³⁸

Anarchy and War

We have seen that offensive Realists believe that the anarchic nature of international politics forces states to be aggressive, while defensive Realists believe that even states inclined to aggression are forced by the anarchic structure of the system to create balances of power instead. All structural Realists agree, however, that interstate wars will continue to occur as long as there is no world government.

Some advocates of world government would agree with this proposition, and their only disagreement with structural Realists concerns the fea-

36. For one of the few systematic discussions of how the term might be defined, see Levy 1983, 10–19.

37. For further discussion, see Wagner 1993. On the importance of third states for understanding the cold war, see especially Trachtenberg's (1999) discussion of the role of Germany.

38. Schweller, in an analysis of what he calls "tripolarity," defines a "pole" as a state that possesses "at least half of the resources of the most powerful state in the system" (1998, 46). But he then proceeds to discuss tripolar systems as though they were three-state systems and therefore, like Waltz, confuses the number of states with the distribution of power among them. Moreover, Schweller includes forces in being as part of his measure of military capabilities. But forces in being are a function of the decisions made by states and therefore cannot be part of the structure of an international system, which is supposed to constrain the decisions of states.

sibility and/or desirability of world government. Many other people have thought, however, that peace could be achieved without world government. Trade, democracy, socialism, international institutions that fall short of being a world government, or just a common realization that war is self-defeating have all been advanced as possible causes of peace among states. Structural Realists are pessimistic about these suggestions, not because they have examined each of them and concluded it would not have the predicted effects (though there are, of course, many disagreements about the predicted effects of all these factors) but because they think they have an argument that shows that none of these factors, *or any others that one might suggest*, could possibly eliminate interstate wars. This is a very strong claim. Is there any reason to believe it is true?

The claim is that in any anarchic system wars will occur. *Anarchy* just means that there is no world government, so the implicit premises in the structural Realist argument include at least some of those stated by Mearsheimer:

Premise 1 (Anarchy) *There is no world government.*

Premise 2 (Anarchy) *All states are capable of using force against other states.*

Premise 3 (Anarchy) *All states seek to maintain their territorial integrity and domestic autonomy.*

The justification for thinking that this last premise is part of what is meant by anarchy is, perhaps, that if it were not true states would abandon anarchy and create a world government.

In addition, since our goal is to establish what is the best we could expect from an anarchic system, it makes sense to use another of Mearsheimer's premises:

Premise 4 (Best case assumption) *States are rational actors.*

It is not really clear what *rational* means here, but a case for the possibility of peace that relied on the irrationality of states would not be a very strong case, so let us stick it in here and worry about exactly what it means later.

This appears to me to be about as far as one can go in writing down what is implied simply by saying that the interstate system is anarchic. But what are these premises together supposed to imply about the occurrence of war? Waltz has virtually nothing to say about that in *Theory of International Politics* (1979), which is usually taken to be the canonical statement of structural Realism, and what he does say seems to contradict the claim that anarchy is an important part of the explanation of war. The main claim in that book is, rather, that anarchy leads to balances of power, as

we have seen. To see what Waltz thought about the relation between anarchy and war, one must look at his first book, *Man, the State and War* (1959), and a later article entitled “The Origins of War in Neorealist Theory” (1988).

Many people seem to have the impression that Waltz’s structural Realism is Hobbes’s account of the state of nature in modern dress. But Waltz’s inspiration was Rousseau, not Hobbes, and in Waltz’s first book he attributes to Rousseau the idea that

wars occur because there is nothing to prevent them. Rousseau’s analysis explains the recurrence of war without explaining any given war. He tells us that war may at any moment occur, and he tells us why this is so. (1959, 232)

The conclusion Waltz expects us to derive from premises that describe an anarchic interstate system, therefore, would appear to be this one:

Conclusion 1 (Structural Realism) *War may at any moment occur.*

But this conclusion plainly does not follow from these premises.

What is missing? The premises obviously imply that in anarchy, as Waltz said, there is nothing to prevent states from using force if they want to, but the conclusion would be true only if at any moment some state may want to. Suppose, then, we added a premise that said, “At any moment some state may want to use force against another state.” But this would beg the question! No one would doubt that states can use force whenever they want to. The question we started with was whether trade or democracy or something else might lead them not to want to do so. Structural Realism claims that anarchy makes this impossible, but it turns out that Waltz just assumed that it is. The reason this is not obvious is that Waltz confuses the *possibility* of war, which cannot be doubted, with its *probability*, which is what is in question.³⁹

Lest one think I am being unfair to Waltz, the following passage makes crystal clear that that is exactly what he did:

According to the third image [i.e., structural Realism], there is a constant possibility of war in a world in which there are two or more states each seeking to promote a set of interests and having no agency above them upon which they can rely for protection. But many liberals and socialist revisionists deny . . . the possibil-

39. As long as planes fly, plane crashes will be possible. But they are sufficiently infrequent that most people do not worry about them.

ity that wars would occur in a world of political or social democracies. An understanding of the third image makes clear that the expectation would be justified only if the minimum interest of states in preserving themselves became the maximum interest of all of them—and each could rely fully upon the steadfast adherence to this definition by all of the others. Stating the condition makes apparent the utopian quality of liberal and socialist expectations. (Waltz 1959, 227)

The expectations may possibly be utopian, but if they are, it is plainly not the absence of a world government that makes them so.

In the passage just quoted, Waltz said that for an anarchic system to be peaceful it would be necessary not only that no state wanted to use force but also that all states knew that this was true. To many structural Realists this is the key to the relation between anarchy and war. It seems to suggest the possibility of a much stronger and more interesting conclusion than the one Waltz got from Rousseau:

Conclusion 2 (Structural Realism) *At any moment some state may choose to use force against another state, even if no state expects to gain from doing so.*

Note the similarity of this conclusion to the first of the conclusions that one might attribute to Mearsheimer that was discussed previously. As we have seen, it does not follow from Mearsheimer's premises, and it is not implied by the premises describing a world of independent states just discussed. Some people believe that it is implied by what is called the security dilemma.

The Security Dilemma

Although many people associate the notion of a security dilemma with Waltz's structural Realism, the idea was John Herz's, and it is unclear from Waltz's own writings exactly what he thought about it. It was first presented in an article published in 1950, nine years before Waltz's first book was published. This article defends the same thesis that Waltz defended in his book but offers a different justification for it. Remarkably, Waltz does not even cite it. Herz claimed that anarchy leads to war, not, as Waltz said, because in anarchy there is nothing to prevent it but because

Wherever such anarchic society has existed—and it has existed in most periods of known history on some level—there has arisen

what may be called the “security dilemma” of men, or groups, or their leaders. (Herz 1950, 157)

But what is a security dilemma? According to Herz, groups or individuals who “live alongside each other without being organized into a higher unity”

must be . . . concerned about their security from being attacked, subjected, dominated, or annihilated by other groups and individuals. Striving to attain security from such attack, they are driven to acquire more and more power in order to escape the impact of the power of others. This, in turn, renders the others more insecure and compels them to prepare for the worst. Since none can ever feel entirely secure in such a world of competing units, power competition ensues, and the vicious circle of security and power accumulation is on. (157)

He goes on to claim that

families and tribes may overcome the power game in their internal relations in order to face other families or tribes; larger groups may overcome it to face other classes unitedly; entire nations may compose their internal conflicts in order to face other nations. But ultimately, somewhere, conflicts caused by the security dilemma are bound to emerge among political units of power. (158)

It appears that Herz was an “offensive Realist” long before Mearsheimer was. We have already seen that Waltz explicitly denied that states seeking security from attack will, as Herz claimed, be “driven to acquire more and more power in order to escape the impact of the power of others.” However, in his later article “The Origins of War in Neorealist Theory,” Waltz cites Herz’s article and says that the security dilemma is the link between anarchy and war (1988, 619). This adds to the puzzle of what Waltz thought the connection between anarchy and war really was.

Whatever the answer to that question might be, the more important question is whether (1) there is any reason to think that anarchy must lead to a security dilemma and (2) there is any reason to think that security dilemmas lead to war. Let us begin by looking at what Herz says a security dilemma is.

The most common interpretation of what Herz had in mind seems also to be the most plausible one. It can be captured in two premises, which we can add to the ones listed previously that describe a world of independent states. The first premise is as follows:

Premise 5 (Security dilemma) *An increase in one state's ability to protect itself from an attack by others will diminish the ability of other states to protect themselves from an attack by the first state.*

This seems to be an obvious consequence of the fact that military power is relative, so that, for example, the number of infantry divisions one state needs to defend itself against another depends on how many infantry divisions the other state has, and so an increase in the size of one state's military forces will reduce the chances of success of its potential adversary.

However, no state would care about the size of another state's military forces if it were certain that the other state would never use them in an attack. And therefore if the fact that power is relative is to have any significance we need the following premise as well:

Premise 6 (Security dilemma) *No state can ever be certain that another state will not use force against it.*

Note that this is identical to one of Mearsheimer's premises discussed earlier.

In fact, we now have a set of premises identical to Mearsheimer's except for the addition of one premise, which merely states the obvious fact that only relative, not absolute, military capabilities are important for a state's security. It is obvious that neither conclusion 1 nor conclusion 2 follows from these premises, any more than the various possible interpretations of Mearsheimer's conclusions followed from his. Thus the security dilemma is of no help whatsoever in showing that it is utopian to think that a world of independent states might be peaceful.⁴⁰

Before moving on, let us notice exactly what is missing in this argument. It is certainly possible that, in a world of independent states, no state would actually expect to benefit from war. Thus if anarchy alone is to lead to war, then war must be possible even if no state expects to benefit from it. But if no state actually uses force against another, then no war will occur no matter how apprehensive states might be about its possibility. Thus if anarchy alone is to lead to war, there must be some reason to expect that anarchy alone will lead a state to use force against another state *merely because it fears that another state might use force against it*. There are obviously occasions when states might do such a thing. Indeed, that is what the United States and Britain did against Iraq in the second war in the Persian Gulf. However, the conditions that make that possible

40. As noted earlier, it is far from clear why we should believe that premise 6 is true. The important point to note here, however, is that, even if it is true, structural Realism's main claim about the relation between anarchy and war is still not supported.

are not implied simply by the absence of a world government. Nor, as we will see, would a world government make it impossible for such conditions to exist.

Offense and Defense

In addition to doubting whether the sixth premise is true, one might also doubt whether the fifth premise must be true. Clearly it is true of infantry and tanks. But it does not seem to be true of fortifications: increasing the strength or number of one state's castles does not appear to diminish the effectiveness of another state's castles.⁴¹ Moreover, the condition that came to be known as mutual assured destruction (or MAD) seems to imply that it is not true of nuclear weapons either: once a state with nuclear weapons has a secure second strike capability, it is not obvious that it needs more nuclear weapons or that, if it acquires more, its adversary's ability to protect its independence is diminished.

Examples such as these led Robert Jervis to argue, in one of the most influential articles about international politics ever written, that the truth of the fifth premise depended on two factors: "whether defensive weapons and policies can be distinguished from offensive ones, and whether the defense or the offense has the advantage" (Jervis 1978, 186).⁴² Paradoxically, even though Jervis questioned whether the fifth premise was always true, his article nonetheless helped convince many people that the security dilemma was the key to understanding why war occurred. The reason is that his article seemed to explain something that was otherwise unexplained by structural Realism.

Even people who did not ask whether Waltz's structural Realist argument was valid noticed that anarchy was a constant property of international politics, but the frequency of war varied greatly. The only explanation of that fact that Waltz had to offer was the distinction between bipolar and multipolar systems. But he claimed that

Until 1945 the nation-state system was multipolar, and always with two or more powers. In all of modern history the structure of

41. Be careful not to be misled by this example, however: castles are fairly immobile, but fortifications can be constructed as infantry forces advance, and so they are more mobile than they might appear to be. If they can be used to protect troops as they advance against the enemy, then they add to the effectiveness of infantry.

42. The effect of the relation between offensive and defensive capabilities on the likelihood of war was the subject of a book by George Quester (1977) published shortly before Jervis's famous article. Jervis's article helped stimulate a long debate about this question that is still going on. Representative selections from the literature can be found in Brown et al. 2004. For a recent survey, see Morrow 2002.

international politics has changed only once. We have only two systems to observe. (Waltz 1979, 163)

Yet prior to the “long peace” of the cold war, there was the “long peace” of the nineteenth century. The distinction between bipolarity and multipolarity could not account for this. Perhaps what Jervis called the “offense-defense balance” could.

Thus Jervis’s discussion of the offense-defense balance only reinforced many people’s belief that the most fundamental cause of war is indeed anarchy and the security dilemma, though the security dilemma may be attenuated if military technology favors the defense. There are, however, two problems with this inference. First, as we have already seen, anarchy and the security dilemma alone do not imply that war will occur, even if defensive capabilities are not dominant. Thus anarchy and the security dilemma cannot explain why wars occur even when the conditions described by the security dilemma exist. And therefore, second, there is no reason to believe that changing the offense-defense balance will change the likelihood of war at all.

We have already noted that one of the things missing in Waltz’s discussion of the relation between polarity and war is any explanation of why wars ever occur. The same is true of both Herz’s and Jervis’s discussion of the security dilemma. Without such an explanation it is not possible to say what the effect of the offense-defense balance might be on the likelihood of war. This question will be the focus of subsequent chapters.

Even without such an explanation, however, there are good reasons to doubt Jervis’s claim that changing the offense-defense balance will change the likelihood of war. For one effect of a situation in which the defense has the advantage might be that states that fight each other are unlikely to risk being disarmed. If so, then war would be less risky than if the offense had the advantage and therefore possibly more attractive.

It is unlikely, for example, that the Palestinians thought in recent years that they could defeat the Israeli army. At the same time Israel has been unable to disarm the Palestinians. Surely this does not make the armed conflict between them puzzling but instead helps explain why it occurs—otherwise the Palestinians either would never have dared challenge the Israelis or would long ago have been disarmed by them. Similarly, the eighteenth century was characterized by frequent wars, and the nineteenth by a long peace. Yet it seems odd to suggest that the offense had a greater advantage in the eighteenth century than in the nineteenth—in fact, exactly the opposite seems to be true. It seems unlikely, therefore, that variations in the salience of the security dilemma could actually account for the difference between them.⁴³

43. For a development of this point, see Fearon 1995a.

The Security Dilemma and the Prisoner's Dilemma

In his famous article, Jervis also gave another reason to question whether the security dilemma must have dire consequences. He claimed that it could be represented by the famous 2×2 game commonly called the Prisoner's Dilemma, in which "there is no solution that is in the best interests of all the participants." This would seem to justify pessimism about the ability of independent states to avoid conflict. However, if the Prisoner's Dilemma game is expected to be repeated indefinitely, he said, then cooperation becomes possible, though still not certain. This would imply that violent conflict among states might be avoidable (Jervis 1978, 171).

By offering the Prisoner's Dilemma game as a model for the security dilemma, Jervis seemed to provide additional support for the view that the security dilemma was the key to understanding the recurrence of war among independent states. At the same time, by showing that the security dilemma did not make war inevitable, this model provided yet another way of showing how anarchy, which was constant, could explain the occurrence of war, which was not.

Unfortunately, the only connection between the security dilemma and the Prisoner's Dilemma is that they both have the word *dilemma* in their names. And therefore, like his discussion of the offense-defense balance, the additional plausibility that Jervis's use of the Prisoner's Dilemma gave to the idea that anarchy made peace among independent states unlikely was quite unwarranted.

The security dilemma, remember, is represented by premises 5 and 6. The Prisoner's Dilemma is represented in figure 1.⁴⁴ The "dilemma" in the Prisoner's Dilemma is the result of two facts, both obvious from figure 1: (1) each player would want to choose D, whatever he expects the other to do, but (2) if they do that, they will end up with an outcome worse for both than if they had both chosen C.

	C	D
C	(3, 3)	(1, 4)
D	(4, 1)	(2, 2)

Fig. 1. The Prisoner's Dilemma

In thinking about what relation there could possibly be between these two so-called dilemmas, note first that part of the definition of the security

44. The first number in each cell is the payoff to the row player, and the second is the payoff to the column player. They merely represent the preferences of the players, with bigger numbers being preferred to smaller ones. The labeling of the choices is conventional and is the result of the fact that they are customarily thought of as "Cooperate" or "Defect." However, the names have no significance whatever.

dilemma is that states are uncertain about what each other's preferences actually are. However, in the Prisoner's Dilemma it is assumed that the players' preferences are commonly known and there is no uncertainty about them. Note second that in the Prisoner's Dilemma the essence of the problem is not the players' preferences but the constraints under which they must choose: they must *independently* choose only once between two alternatives. The security dilemma, however, does not specify what choices states must make or how they will go about making them—that is why one cannot infer from it anything about what choices they can be expected to make. It merely says that, when they choose what military capabilities to have, an increase or decrease in one state's military capabilities will change the relative size of the other's.

Suppose, for example, that states are choosing between arming and not arming. Then, one might think, premise 5 would imply that they would have the preferences represented in *Figure 1*. But if states can observe each other's arms levels and respond to them, their choices would not be restricted to simply arming or not but would include as well the possibility of arming if the other does but not arming if it does not. That would not be a Prisoner's Dilemma, even if states had the preferences represented in *Figure 1*. Moreover, if they did both arm they would be worse off than if they had not, but that does not imply that they would *fight* each other.⁴⁵

Of course, if the Prisoner's Dilemma is expected to be repeated, then the choice that a state makes in one round can be based on the choice that the other state made in the previous round. But each must still pay the price of the other's defection if that is what it chose. It is not clear why this must be so or even what it would mean in any concrete case. Moreover, if the choice involves war (as it must, if the security dilemma is to explain why wars occur), then it is not clear why states would expect that exactly the same game would be played again after a war occurred. In fact, one might think that if states could expect that they would always be around to play the next stage of a repeated Prisoner's Dilemma game, they would enjoy a far greater degree of security than Herz thought they could possibly have.

Finally, let us note that Jervis claimed not only that the security dilemma could be represented by a repeated Prisoner's Dilemma but also that a repeated Prisoner's Dilemma could be represented by the 2×2 game often called the Stag Hunt. The Stag Hunt is portrayed in *Figure 2*. Note

45. Jervis says, "A relatively low cost of CD has the effect of transforming the game from one in which both players make their choices simultaneously to one in which an actor can make his choice after the other has moved" (1978, 172). This is obviously not true. But there is no reason to assume that states must make their choices simultaneously in the *first* place.

that it is different from the Prisoner’s Dilemma only in that the players’ preferences between joint cooperation and defecting when the other cooperates are reversed. Now, if one player expects the other to choose C, he will want to choose C as well, whereas each will want to choose D only if he expects the other to choose D. In the lingo of game theory, this means that, instead of each having a dominant choice of D, there are two pure strategy Nash equilibria, CC and DD, and determining which will occur depends on what each expects the other to do.⁴⁶

	C	D
C	(4, 4)	(1, 3)
D	(3, 1)	(2, 2)

Fig. 2. The Stag Hunt

However, while in this game, as in the repeated Prisoner’s Dilemma, both joint cooperation and joint defection are equilibria, the Stag Hunt is not the repeated Prisoner’s Dilemma but is instead just another one-shot 2×2 game like the one-shot Prisoner’s Dilemma. Furthermore, in the Stag Hunt the preferences of the players are common knowledge, and therefore the uncertainty confronted by the players of this game, if there is any, is the result of the existence of more than one equilibrium in the game and not uncertainty about what the other state’s preferences might be. But the security dilemma is defined by uncertainty about what other states’ preferences might be.

The name “Prisoner’s Dilemma” is based on the fact that the game was originally illustrated by a story about two prisoners who were induced to confess to a crime by a clever district attorney who, by separating them, forced them to choose independently between confessing or not confessing. The name “Stag Hunt” is based on the fact that the game with that name seems to represent a situation described by Rousseau in his *Discourse on the Origins of Inequality*, which Waltz had used in his book *Man, the State and War* to illustrate the effect of anarchy (Waltz 1959, 167–71).

However, Rousseau used the story not to make a point about international politics but to illustrate what he thought was the lack of foresight of primitive men, who “were so far from troubling themselves about the distant future, that they hardly thought of the morrow.”

46. A Nash equilibrium is just a set of plans for making choices (called a “strategy”), one for each individual, such that if everyone expects everyone else to choose the appropriate strategy in this set, no one would have any incentive to choose some other strategy. There is also a mixed strategy equilibrium in this game, but I will ignore it.

If a deer was to be taken, every one saw that, in order to succeed, he must abide faithfully by his post: but if a hare happened to come within the reach of any one of them, it is not to be doubted that he pursued it without scruple, and, having seized his prey, cared very little, if by so doing he caused his companions to miss theirs. (Rousseau 1913, 194)

Rousseau thought that the ugly nature of international politics was the indirect result of the fact that human beings had learned only too well to be more industrious, farsighted, and cooperative than that. And Waltz used the story merely to illustrate the proposition that what is rational for an individual is not always rational for a group (1959, 168–71).⁴⁷

Nonetheless, Waltz's use of Rousseau's story and Jervis's subsequent use of the 2×2 game with the same name have led many people to believe that the Stag Hunt game contains an important insight about the nature of international politics. To add to the confusion, in his *Theory of International Politics*, which was published one year later, Waltz claimed that in international politics "states face a 'prisoner's dilemma'" and cited Jervis's article in support of that claim (Waltz 1979, 109).

Jervis's use of these famous games reflected a more general tendency at the time to think that the family of 2×2 games, each of which differs from the others only by having a different configuration of preferences, were ready-made models of any social situation of interest.⁴⁸ However, as we have seen, these games contain hidden assumptions that are very strong. Many people thought their use could be justified by the fact that models are not supposed to be descriptively accurate, which is true. However, the assumptions represented by 2×2 games are frequently inconsistent with other assumptions made by people who use them to justify their conclusions, and therefore the arguments they support are self-contradictory.

What 2×2 games can do is serve as *examples* of the counterintuitive effects of the interdependence of choices that can also occur in more complex social situations. The Prisoner's Dilemma illustrates the fact that mutually beneficial choices may not be made if individuals have an incentive to take advantage of other people's decisions to cooperate. The Stag Hunt illustrates the fact that even if this problem does not exist cooperation may not occur if people are not sufficiently confident that others intend to cooperate as well. These are certainly problems that arise in international politics, but they are not restricted to international politics.

47. For a recent development of the idea that the Stag Hunt game, and not the Prisoner's Dilemma, best represents the fundamental problem of human social organization, see Skyrms 2004.

48. For an especially influential example, see Snyder and Diesing 1977.

And it is not obvious what they have to do with the security dilemma defined by Herz.

These early attempts to use game theory as a way of thinking about the security dilemma were hampered by the fact that, at the time, game theory provided no way of thinking about one of the defining features of the security dilemma, the uncertainty of states about other states' preferences. Moreover, equilibrium outcomes in matrix representations of games implied implausible predictions when one looked at the actual sequence of choices represented by a game tree. Subsequent developments in game theory that addressed both these issues have led to the widespread use of game models in extensive form with incomplete information. Ironically, unlike the earlier misleading use of 2×2 games, they are the cause of complaints that trivial models have driven out significant research in political science. But they have only focused attention on all the complexities that models like the Prisoner's Dilemma and the Stag Hunt concealed.⁴⁹

Hierarchy and Peace

If anarchy is the root cause of war, one would expect that government (or "hierarchy," as Waltz called it) should lead to peace. But in his famous book *Theory of International Politics*, Waltz explicitly and emphatically denied that this was true:

The threat of violence and the recurrent use of force are said to distinguish international from national affairs. But in the history of the world surely most rulers have had to bear in mind that their subjects might use force to resist or overthrow them. If the absence of government is associated with the threat of violence, so also is its presence. The most destructive wars of the hundred years following the defeat of Napoleon took place not among states but *within* them. . . . If the possible and the actual use of force mark both national and international orders, then no durable distinction between the two realms can be drawn in terms of the use or the nonuse of force. No human order is proof against violence. (1979, 102–3)

This passage seems to contradict the main thesis of *Man, the State and War*. Nonetheless, Waltz later wrote:

49. For a discussion of what can be learned by analyzing the implications of assuming that participants in a 2×2 game are uncertain whether the other person's preferences conform to the preferences in the Prisoner's Dilemma or to those in the Stag Hunt, see Kydd 2005.

Although neorealist theory does not explain why particular wars are fought, it does explain war's dismal recurrence through the millennia. . . . The origins of hot wars lie in cold wars, and the origins of cold wars are found in the anarchic ordering of the international arena.

The recurrence of war is explained by the structure of the international system. (1988, 620)

But what, then, explains the recurrence of civil wars?

These apparent contradictions are in part yet another illustration of the fact that structural Realism has virtually nothing to say about why war ever occurs anywhere. It is therefore not surprising that structural Realists actually have nothing to say about the connection between either anarchy or hierarchy and the occurrence of war. But these passages also illustrate another important fact about structural Realism: in spite of the fact that its main theme is the difference between anarchic and hierarchical systems, it also has little to say about what that difference is.

Here is what Waltz had to say about it:

The difference between national and international politics lies not in the use of force but in the different modes of organization for doing something about it. . . . A government has no monopoly on the use of force as is all too evident. An effective government, however, has a monopoly on the *legitimate* use of force, and legitimate here means that public agents are organized to prevent and to counter the private use of force. Citizens need not prepare to defend themselves. Public agencies do that. A national system is not one of self-help. The international system is. (1979, 103–4)

But the world is full of governments that are not “effective” in this sense, and yet neither Waltz nor any of his followers have ever suggested that structural Realism might have something to say about them. Moreover, the origins of the U.S. Civil War are not to be found in the fact that in the nineteenth century U.S. citizens could not look to government to protect them against the private use of force.

The end of the cold war has been followed by a period like the one in the nineteenth century described by Waltz, in which the most destructive wars have taken place “not among states but within them.” As one might expect from these passages in which Waltz attempted to state the difference between anarchy and hierarchy, structural Realism has been little help in understanding them.

For example, one Realist, Barry Posen, has written that they can be understood as the result of the collapse of governments: when the sover-

eign disappears ethnic groups are faced with the security dilemma that results from anarchy (1993). John Mearsheimer, however, argued that only partition could resolve the conflict in Kosovo or provide a long-term resolution of the conflicts in Croatia and Bosnia (1998). But even if the security dilemma could explain why war occurs (which, as we have seen, it cannot), it would seem strange to say that in 1860 the U.S. government first collapsed and then the North fought the South because of the resulting security dilemma. And the consequence of partition is to substitute anarchy for a common government. If anarchy has the consequences Mearsheimer claimed for it, how could it lead to peace among warring ethnic groups?⁵⁰

Compare these conflicts with the recent conflicts between India and Pakistan. India and Pakistan are independent states, each with nuclear weapons, and the conflicts between them might be taken to illustrate the dire effects of anarchy and the security dilemma. However, their existence is the consequence of the partition of British India into two states, one predominantly Hindu and the other predominantly Muslim, after it became independent. Would conflicts between Hindus and Muslims in the Indian subcontinent be greater or less if India had not been partitioned? Since the differential effect of anarchy and hierarchy on violent conflict is one of the main themes of structural Realism, one might expect that it would have an answer to that question. But in fact it does not.

In this respect structural Realists are not unique. Virtually everyone takes governments for granted, and this fact is reflected in the division of intellectual labor among American political scientists between students of international politics and students of comparative or domestic politics: international politics is the study of relations among governments, and everything else is the study of politics structured by governments within borders that define their jurisdictions. But governments and borders come and go, the incidence of organized violence within the jurisdictions of governments varies enormously across time and space, and both well-defined borders and governments that resemble the ones in contemporary Europe or the United States are rare and a very recent development. The way political scientists organize their work has created a gap in our knowledge about the central problem of our time: the relation between political institutions and organized violence. Structural Realism is like a rug thrown over this gap that only makes it harder to see it.

This gap in our knowledge not only inhibits our understanding of what is now called “state failure.” It also inhibits our understanding of international institutions. In discussing the idea of world government, Inis

50. On this question, see also the section of Waltz's *Theory of International Politics* called “The Virtues of Anarchy” (1979, 111–14).

Claude said:

In the final analysis, it appears that the theory of world government does not *answer* the question of how the world can be saved from catastrophic international conflict. Rather, it helps us to *restate* the question: How can the world achieve the degree of assurance that inter-group conflicts will be resolved or contained by political rather than violent means that has been achieved in the most effectively governed states? (1962, 271)

Because of the division of intellectual labor between students of domestic and international politics, we still do not have a good answer to Claude's question. And therefore we do not know what contribution international institutions short of a world government might make to the resolution of interstate conflicts. Structural Realism does not answer this question, it begs it—as do the advocates of various interstate institutions like the International Criminal Court, whose main appeal is that they resemble some of the features of modern states.

Worse than that, we cannot even specify clearly what the difference between government and anarchy is. Is the European Union a government, and if not, what would suffice to turn it into one? Is there a government in Yemen? Was there a government in Afghanistan under the Taliban? Does an Afghan warlord preside over a government? Does the rebel group in Colombia known as the FARC, which controls a large segment of the territory nominally allocated to the government in Bogotá, constitute a government in the territory that it controls? Was there a government of the United States prior to the U.S. Civil War? How long has there been a government of France, and when did it first appear? The question that structural Realism begs is not even well defined.⁵¹

Realism's Competitors

Structural Realism makes three main claims: (1) the anarchic structure of international politics leads to the recurrence of war, (2) war is less frequent in anarchic systems with a bipolar structure than in systems with a multipolar structure, and (3) in anarchic systems with a multipolar structure, alliances lead to balances of power rather than to a preponderance of

51. For an elaboration of this point, see Milner 1991. See also Lake 2003. A good way to begin thinking about this question is to immerse oneself in Samuel Finer's great posthumous work, *The History of Government* (1997). For a survey of the anthropological literature on the development of the state, see Johnson and Earle 1987. An older discussion by an anthropologist that I have found very helpful is Fried 1967.

power. No valid argument has ever been presented in support of any of those claims, which is what makes the disagreement between defensive Realists and offensive Realists possible. Thus the only good reason for continued attention to the works that defined structural Realism is that they provide illustrations of how easy it is to make mistakes in thinking about international politics.

However, in a volume recently published under the auspices of the American Political Science Association describing the current “state of the discipline,” Stephen Walt claims that “[t]he bottom line is that realist theory is alive and well. It remains relevant, rigorous, and theoretically fecund” (2002, 222). In support of his evaluation, Walt says that the “utility of a research tradition may be judged by two basic criteria,” which he calls “explanatory power” and “internal fertility.” Following Van Evera (1997), he says that “explanatory power can be judged by the percentage of variance explained by the independent variable(s), the range of topics covered by the theory, and the prevalence of the phenomena being explained” (Walt 2002, 201). And therefore, he claims that

debates within the realist family and between supporters of realist theory and those of various rivals should not be resolved by asking who can muster the shiest abstract argument; rather we should ask which explanation best ts the facts. Determining which theory (or approach) is most useful is an empirical question, and rendering such judgments usually requires careful historical evaluation of the specific causal mechanisms in each theory. (Walt 2002, 224)

These statements illustrate the fact that it never occurs to many political scientists that logical validity is an important criterion in evaluating arguments. Empirical evidence cannot confirm or disconfirm an explanation if the evidence is not actually implied by the explanation. For that, the argument need not be “shy,” but it must be valid.⁵²

As these quotations from Walt illustrate, one reason for the low value political scientists place on logical validity is the widespread confusion among them between explanations and regression equations. Another is the tendency to equate a theory with any plausible conjecture. Various buzzwords from the philosophy of science are often invoked as ways of

52. For Van Evera’s own use of the criteria for evaluating theories listed by Walt, see Van Evera 1999. In an earlier article attacking the use of formal models, Walt (1999) claimed that it was consistency that we should aim for in constructing explanations. But the problem with structural Realism is not that it is inconsistent but that the arguments offered in support of its main claims are not valid. For another influential example of this confusion, see King, Keohane, and Verba 1994, 105–7.

evaluating those conjectures. But ultimately many political scientists believe that it is the job of empirical research to resolve the disagreements among the authors of these conjectures.⁵³

This can never happen, because no facts can actually be derived from any of the competing “theories.” Thus competing conjectures accumulate. Similar conjectures are grouped into schools of thought and named, and scholars who find them plausible identify themselves with them (as Walt identified himself as a member of the “realist family”). Students in political science courses are then expected to know about these competing families, and if the question of how to evaluate them arises, it is only by asking students to think about how they might be tested empirically. And therefore great academic rewards go to those who succeed in devising a new “paradigm” or “approach,” which can then provide the basis for further inconclusive empirical research. Thus, in spite of all the talk about science and scientific method among political scientists, the study of international politics does not satisfy the definition of a science given at the beginning of this chapter.

This is not only sad but ironic, since Waltz’s book *Theory of International Politics* was written to counter just such attributes of the field. “Among the depressing features of international-political studies,” he wrote,

is the small gain in explanatory power that has come from the large amount of work done in recent decades. Nothing seems to accumulate, not even criticism. Instead, the same sorts of summary and superficial criticisms are made over and over again, and the same sorts of errors are repeated. (Waltz 1979, 18)

The first chapter of that book is devoted to emphasizing the distinction between a correlation and an explanation, and in it Waltz emphasized that an explanation required a creative guess as to what propositions might imply the facts to be explained. Unfortunately, Waltz’s arguments did not satisfy his own criteria, but most people accepted his claim that they did. As a result, many criticisms of Waltz’s ideas are criticisms of what some people take to be his assumptions rather than the validity of his arguments.

One reason many people accepted Waltz’s arguments so uncritically is probably that they already believed his main conclusions to be true. The

53. One favorite criterion for evaluating theories is “parsimony.” Note that it is hard to beat a bald general assertion for parsimony. See the recent collection of essays in Elman and Elman 2003, which try to appraise the status of international relations theory by the standards laid down by Imre Lakatos but in which no one asks the simple but obvious question whether the arguments being evaluated are valid. See also James 2002.

idea that states form coalitions to balance the power of other states is one of the oldest ideas in writings about international politics. During the cold war the United States and the USSR were obviously much more powerful than other states, and it seemed very plausible that that fact explained both the conflict between them and their ability to avoid war with each other. And the proposition that violent conflicts were likely without a government to prevent them also seemed obviously true to many people.

Waltz's emphasis on the anarchic nature of the international system seemed especially compelling. It implicitly invoked not only the support of the entire Western tradition of the social contract and its concept of the state of nature but also the theory of the state that had been developed by economists, in which the function of the state was to supply public goods and compensate for market imperfections. And, as the article by Jervis discussed previously illustrated, the Prisoner's Dilemma game had come to be accepted by many people as a persuasive illustration of why government was necessary if people's common interests were to be served and therefore a validation of pessimistic expectations about the consequences of anarchy or the state of nature (Jervis 1978).⁵⁴

Neoliberal Institutionalism

The works of two economists in particular were especially influential in reinforcing the view that the anarchic nature of international politics was the key to understanding what happened in it: Albert Hirschman and Charles Kindleberger. In a book about the prospects for constructing a peaceful international order after World War II, Hirschman pointed out that sovereign states always had the option of interrupting trade with other states, in which case the gains from trade became the losses from the interruption of trade. If those losses were not distributed symmetrically, he claimed, they could be the basis for demanding political concessions, and therefore international trade necessarily had an impact on the ability of sovereign states to exercise influence over each other (Hirschman 1945). And in a book about the Great Depression, Charles Kindleberger argued that, given the lack of an international monetary authority, international monetary stability required a dominant state willing to act as a substitute. Britain, he claimed, had served that function in the nineteenth century; in

54. It is ironic that Ruggie has criticized the use of rational choice models to study international politics because they resemble the models found in economics, which he claims (oddly) require the existence of markets; and yet, he claims, they cannot explain where markets come from (1998, 23). But economists have traditionally thought that *governments* were required to make markets work, which makes structural Realism's claim that bad things should be expected in a condition of anarchy seem plausible—a claim that Ruggie wants to dispute.

the twentieth century the only substitute available was the United States, but in the 1930s it was unwilling to assume that role (Kindleberger 1973).

Kindleberger's thesis was one of the bases of what is known as "hegemonic stability theory," which came to be accepted as part of Realism.⁵⁵ But the only connection between hegemonic stability theory and structural Realism was that they both purported to be consequences of the anarchic nature of international politics (Keohane 1984, 7–10).

As the economic preponderance of the United States declined, hegemonic stability theory seemed to imply that people should be pessimistic about the future of the international economy. But in an influential book called *After Hegemony*, Robert Keohane (1984) argued that such a pessimistic conclusion did not follow from anarchy alone. Rather, he claimed, independent states could cooperate in an anarchic environment, and international institutions could facilitate such cooperation. The repeated Prisoner's Dilemma was one of the foundations of his argument (65–84). This thesis came to be known as "neoliberal institutionalism," and another "ism" was added to the world's inventory of doctrines.

Keohane claimed to have shown that

even on the restrictive assumptions of Realism and game theory, gloomy conclusions about the inevitability of discord and the impossibility of cooperation do not logically follow. Egoistic governments can rationally seek to form international regimes on the basis of shared interests. (107)

The first statement is certainly correct—in fact, none of the main claims made by structural Realists follows from their assumptions. It is therefore unclear why, in the conclusion to his book, Keohane praised the "taut logical structure" of Realism and said that it "should not be discarded, since its insights are fundamental to an understanding of world politics," but that "it does need to be reformulated" (245).

In fact, because of structural Realism's lack of a "taut logical structure," it is not even clear that Keohane's conclusions are inconsistent with what Waltz said about the consequences of anarchy. Keohane claimed that

If there were an infinitely large number of equally small actors in world politics . . . [i]nternational conditions would more closely approximate the Hobbesian model in which life is "nasty, brutish, and short." But as we have seen, the fact that the number of key actors in the international political economy of the advanced

55. See also Gilpin 1975, 1981.

industrial countries is typically small gives each state incentives to make and keep commitments so that others may be persuaded to do so. (258)

It is hard to see how this contradicts Waltz's discussion of the benefits of the fact that economic interdependence among states is managed by a small number of great powers.⁵⁶

Nonetheless, members of the Realist family resisted Keohane's attempt to hoist them by their own petard, the Prisoner's Dilemma. Prominent among them was Joseph Grieco, who seems to have coined the term *neoliberal institutionalism* (1993, 335–36). Grieco (1988) claimed that neoliberal institutionalists had overlooked the fact that states in an anarchic environment had to be concerned not just about achieving gains from economic cooperation but also about protecting themselves from the possibility of elimination through war. But, as we have already noted, the possibility that states might be eliminated through war contradicts the assumptions underlying the repeated Prisoner's Dilemma game.⁵⁷ And without the repeated Prisoner's Dilemma, neither structural Realists nor neoliberal institutionalists have the basis for making any claims whatever about the consequences of anarchy for the behavior of states.⁵⁸

However, Grieco avoided this modest conclusion and claimed instead that Keohane was wrong about states' *preferences*. Neorealists, he claimed, assumed that states were concerned not about *absolute* gains, as in the standard Prisoner's Dilemma game, but about *relative* gains. But structural Realists had not objected to Jervis's use of the standard Prisoner's Dilemma as a model of the security dilemma. Rather, as we have seen, Waltz was happy to accept it as confirmation of his main thesis. Moreover, Grieco did no more to show how a concern for relative gains could be inferred from premises describing anarchy than Waltz or Mearsheimer have done to support their inferences from them.

The resulting "relative gains" controversy helped make the disagreement between neorealists and neoliberal institutionalists a staple of international relations courses all over the country.⁵⁹ Like Walt, Grieco claimed that any resulting disagreements could be resolved by empirical tests.

Neoliberal institutionalism is not based on realist theory; in fact,

56. See especially Waltz 1979, 132.

57. This was pointed out by Robert Powell (1991) in response to Grieco's article.

58. Another strand of Liberal thinking about war and the state rests on the empirical observation that democratic states do not fight each other and takes its theoretical inspiration from Immanuel Kant's response to the writings of Hobbes and Rousseau (Doyle 1983). I will discuss Kant's ideas, and what contemporary writers have made of them, in the next chapter.

59. The main contributions to this controversy are reproduced in Baldwin 1993. For an incisive review and critique of the arguments advanced by the participants in this controversy, see Powell 1994.

realism specifies a wider range of systemic-level constraints on cooperation than does neoliberalism. Thus the next scholarly task is to conduct empirical tests of the two approaches. (Grieco 1988, in Baldwin 1993, 131)

But an empirical test of an “approach” is not possible.⁶⁰

Constructivism

Like neoliberal institutionalism, Constructivism takes as its point of departure the fact that structural Realism’s conclusions cannot be derived from premises that describe anarchy. However, it also objects to the way in which neoliberal institutionalists have tried to think about what the consequences of anarchy might be, as exemplified by the repeated Prisoner’s Dilemma game. In the end, however, it is not really clear what this objection is.

Like structural Realism, Constructivism is more nearly a family of like-minded people than a system of logically related propositions. Two of the most prominent members of this family are John Ruggie and Alexander Wendt. I will focus primarily on Wendt.

“Constructivism,” Wendt says, “is not a theory of international politics” (1999, 7). What, then, is it? Like much else about Constructivism, the answer to that question is not entirely clear.⁶¹ Unlike neoliberal institutionalism, Constructivism’s critique of structural Realism invokes ideas from philosophy and sociology, which are used to support a version of what Dennis Wrong (1961) called “the oversocialized conception of man in modern sociology.”⁶² This conception of man and society is then applied to states and international systems and provides the basis for Wendt’s claim that “anarchy is what states make of it,” that is, that there are no inherent properties of anarchic systems independent of the cultures that define both them and the states they contain—a much stronger claim than Keohane made (Wendt 1992).

Some of what Wendt borrows from philosophy is just the account of inference to the best explanation that I laid out at the very beginning of

60. Note that there is a special problem in looking for evidence of a concern for “relative gains”: whenever states bargain over the terms of a mutually beneficial agreement, no distinction can be made between a concern for absolute gains and a concern for relative gains, since making one state better off must make the other worse off.

61. The term *social constructivism* means somewhat different things in different contexts. My focus here is on what *Constructivism* generally refers to in the literature on international politics. To make this clear I will capitalize the word.

62. Many Constructivist complaints about structural Realism have their origins in Talcott Parsons’s discussion of Hobbes (1937, 43–86). Richard Ashley (1986) was apparently the first to try to interpret structural Realism in this light.

this chapter, which, as I pointed out, is perfectly consistent with Waltz's own account of inductive inference in the first chapter of *Theory of International Politics* (Wendt 1987, 350–55; 1999, 47–91). However, as I also pointed out, we do not need to invoke the authority of philosophers of science to reason in this way, since, as Wendt himself says, “scientists, not philosophers, are the final arbiters of what is ‘scientific’” (1987, 351). Moreover, like Waltz, Wendt overlooks one of the most important implications of this description of inductive inference: if facts are to support explanations, the facts must be implied by the explanation.

The reason Wendt thinks it is necessary to invoke the philosophy of science is that he wants to use this literature to support the proposition that social structures really exist, even though they are not directly observable (Wendt 1987, 351–55). And the reason he thinks this is important is that he wants us to believe that structural Realism rests on the assumption of “methodological individualism,” that is, that only individual people really exist, and therefore statements about social behavior must be reducible to statements about the behavior of individual people. This is a very odd thing to say, since structural Realism is about the behavior of *states*, not individuals, a fact that should make a methodological individualist profoundly suspicious of it.

The relation between individuals and social structures is the subject of what some sociologists call the “agent-structure” problem. To understand it we need only think about the comparison between humans and ants mentioned earlier. Ant social organization is the product of a complex genetic program. Human social organization is invented by humans, but the social organizations humans invent also shape the people who are part of them in complex ways. Every baby is born into a family, and the adult is the product of both the baby's genetic endowment and the family, a fact that gives rise to the nature-nurture controversy among psychologists. Family organization, however, is not simply the product of a genetic program and therefore can be changed. Families, of course, are part of larger cultures and so forth.⁶³

Clearly none of this implies that when individuals encounter each

63. Constructivists like to say that social structures determine the “identities” of individuals. The word *identity* is undefined, and it is not at all clear what it means, especially when applied to states (Fearon 1999). Constructivists writing about international politics, for example, like to say that the cold war ended because the Soviet Union changed its identity. The Soviet Union *lost* its identity when it collapsed, but it is not clear what it means to say that it changed its identity, unless it means simply that people with different ideas became influential or that influential political actors changed their understanding of how the world works. But in that case one might say that every book, including this one, is an attempt to change the identity of its readers. For an extended example of this sort of discussion, see Hopf 2002. Other examples of Constructivist writings about this and other matters can be found in Katzenstein 1996.

other outside of their families or other organizations that they belong to (as, e.g., when they trade with each other), or when families and other groups engage in organized conflict with each other, their interactions do not conform to the assumptions of either the one-shot or repeated Prisoner's Dilemma games (Seabright 2004). Thus not only should all this come as no surprise to any reader, but it is irrelevant to the issues debated by structural Realists and neoliberal institutionalists.

Wendt, however, tries to leave the reader with a different impression. For example, he says:

It is possible for a Hobbesian anarchy to have no culture at all. Here, all knowledge is private rather than shared. . . . The absence of shared culture has an interesting, perhaps counter-intuitive implication: the resulting warfare is not really "war" at all. Killing there may be aplenty, but it is akin to the slaughtering of animals, not war. War is a form of collective intentionality, and as such is only war if both sides *think* it is war. Similarly, a balance of power in this context is not really a "balance of power." Mechanical equilibrium there may be, but actors are not aware of it as such. (Wendt 1999, 266–67)

But this is very misleading. War between organized groups is a "form of collective intentionality" *on the part of each of the warring groups*, but not between them—war is not like a Saturday afternoon game of touch football that people agree to play for their mutual enjoyment. And, unless one simply defines common knowledge as equivalent to a shared culture, whatever level of common knowledge may be required for two groups to form and fight each other, a shared culture, as that word is ordinarily used, is plainly not necessary, either within the two fighting groups or between them.⁶⁴

Wendt says that the subject of his book *Social Theory of International Politics* "was the ontology of international life" (1999, 370). *Ontology* is not a word that any student of international politics should ever have an occasion to use, and therefore it is incumbent on people who use it when talking about international politics to define carefully what they mean by it. Both Ruggie and Wendt use the word freely but never define it. The result can only be to intimidate and confuse the reader.

According to the *American Heritage Dictionary*, ontology is "the branch of metaphysics that deals with the nature of being." It is far from clear how the study of international politics might entangle one in contro-

64. The ability of human beings to infer each other's intentions is not supplied by culture, but it is what makes culture possible (Baron-Cohen 1995; Tomasello 1999).

versies about the nature of being. Wendt says that the question of whether social structures really exist is a question about the “ontological status of unobservables” and therefore similar to the question of whether elementary particles really exist (1987, 351). But to ask whether elementary particles exist is not to ask a question about the nature of being but to ask what evidence there is for their existence, and therefore the question should be addressed to physicists and not philosophers. Similarly, the existence or nonexistence of social structures or shared cultures among human beings who interact with each other is not an ontological question but an empirical one, and the knowledge that they are human beings tells one nothing about the answer to it.⁶⁵

Whatever ontology is, Constructivism’s fundamental objection to structural Realism is that its ontology is (1) individualist and (2) materialist (Wendt 1999, 1–44, 370–78; Ruggie 1998, 1–39). I have just pointed out how strange it is to call structural Realism “individualist.” What could it mean to say that it is “materialist”?

A clue to the answer can be found by looking up the word *materialism* in the dictionary:

materialism n. 1. Philosophy. The theory that physical matter is the only reality and that everything, including thought, feeling, mind, and will, can be explained in terms of matter and physical phenomena. 2. The theory or doctrine that physical well-being and worldly possessions constitute the greatest good and highest value in life. 3. A great or excessive regard for worldly concerns. (*American Heritage Dictionary*)

Social Constructivists complain that structural Realism rests on the assumption that material things like military capabilities, rather than ideas, culture, and social norms, are the only things that influence the behavior of states. This is, perhaps, materialism in the second and third senses of the definition, but it is not clear what it could have to do with something that might plausibly be called ontology. Wendt claims that, since it is materialist, structural Realism emphasizes causality rather than meaning in trying to explain international politics. This might plausibly have something to do with ontology, but it is materialism in the first sense of the definition. Thus Wendt has apparently confused the first meaning of materialism with the second and the third ones. That is why he claims that “[m]aterialists privilege causal relationships, effects, and questions”

65. In the field of artificial intelligence, the word *ontology* refers to a conceptual scheme or system of categories. (A number of ontology editors are available for downloading on the World Wide Web.) To find out more about ontology, see the Buffalo Ontology Web site, at <http://ontology.buffalo.edu/>.

(Wendt 1999, 25) and also claims that theories of bureaucratic politics reflect a materialist orientation (Jepperson, Wendt, and Katzenstein 1996, 33).⁶⁶

Wendt's two complaints about structural Realism can be summarized together by comparing the contrast between ant and human social organization discussed previously, which is summarized in Figure 3, with Wendt's typology of social theories (1999, 29–33). The ant column in Figure 3 exemplifies causal explanations and the human column represents explanations based on meaning and intentionality. An explanation of social behavior that would be genuinely both individualist and materialist (or "physicalist" as Ruggie calls it) would fall into the lower left quadrant. The corresponding explanation of human behavior would, as Weber said, be based on the choices of "participating individual men" and therefore fall into the lower right quadrant. Structural Realism would have to be placed in the upper right quadrant, since it is based on choices made by organizations of human beings acting as units. However, in a table in Wendt's book whose dimensions apparently correspond to the ones in Figure 3, Neorealism (or structural Realism) is placed in the lower left quadrant (Wendt 1999, 32). It is not clear why.

	ant	human
holistic	colonies	state choices
individualist	genetic programs	individual choices

Fig. 3. Ant vs. human social organization

It is clear, however, why Wendt thinks this strange way of characterizing structural Realism is important. He claims that "[n]eorealism's problematic conclusions about international politics" stem not from faulty reasoning but "from its underlying materialist and individualist ontology." This, he claims, implies that "by viewing the system in idealist and holist terms we could arrive at a better understanding" (Wendt 1999, 370).

In other words, if we see (1) that Waltz confused "social kinds" with "natural kinds" (to use Wendt's terminology), we will conclude (2) that the

66. The confusion can be clearly seen throughout Wendt 1999, chap. 3. Actually, it is not clear why structural Realism should be called "materialist" even by the second and third definitions of the term. Cultural or religious differences can lead to military conflicts, and it is not clear what is "materialist" about being concerned about the probability of prevailing in such conflicts. The distinction between "material" and "cultural" explanations is, of course, a staple in debates among sociologists and anthropologists about both Marxism and sociology (Johnson and Earle 1987, 8–11). But it is not clear what any of that has to do with structural Realism. However, the anthropologist Marvin Harris's discussion of what he called "cultural materialism" is directly relevant to many of the issues to be discussed in the following chapters. See, for example, Harris 1977.

nature of international politics is determined by the culture that states share rather than by its anarchy (Wendt 1999, 372). The first statement is hard to square with the fact that Waltz's inspiration was Rousseau, and Wendt considers Rousseau to be a precursor of Constructivism.⁶⁷ But the second would not follow even if it were true.

On this basis, Wendt identifies three types of international culture that might be possible: a Hobbesian culture, a Lockean culture, and a Kantian culture, the main features of each of which are taken from Hobbes, Locke, and Kant respectively (Wendt 1999, 246–312). Using this typology of cultures, he says, for example, that states would form balances of power, as Waltz claimed, only in a Lockean culture in which no state feared elimination by any other state, and therefore a “Lockean culture . . . is a condition of possibility for the truth of Neorealism” (Wendt 1999, 285). But this is simply asserted, and Wendt makes not the slightest attempt to show that it is true or how Waltz was mistaken in thinking otherwise.⁶⁸ Thus Constructivism merely adds yet another collection of invalid arguments with a distinctive name to the field of international politics. Like Liberalism (or neoliberal institutionalism), it is supported indirectly by the inadequacies of structural Realism but fails to identify correctly exactly what those inadequacies are.

In a recent evaluation of Constructivism, Jack Snyder has written that

Current debates about anarchy and culture have been carried out largely at the level of abstract philosophy and visceral morality. Ultimately, however, the impact of culture on war in anarchy is an empirical question. (2002, 9–10)

He then claims that empirical research by anthropologists investigating whether one could “explain behavior in anarchy solely in terms of either cultural or material causes” is relevant to answering this question (12).⁶⁹

In saying this, Snyder illustrates a common view among political scientists that valid inference is just the concern of “abstract philosophy”; any published conjecture is a theory that has to be evaluated empirically; and any argument, valid or invalid, can be represented as a relationship

67. See Wendt 1999, 171. Note that this is the only reference to Rousseau in the index to Wendt's book.

68. If culture alone is sufficient to prevent organized violence at the global level, then one might think that governments are not actually necessary to prevent domestic violence, as structural Realists assume, and wonder why Constructivists are not anarchists. But Wendt does not discuss the problem of domestic order and, indeed, has little to say about organized violence at all.

69. Note that Snyder leaves unclear what the distinction between anarchy and hierarchy refers to in a world of stateless societies or what the connection is between “material explanations” and structural Realists' claims about the effects of anarchy.

between one or more independent variables and a dependent variable. That view will keep Constructivism, along with all the other warring “isms,” in business for a long time to come.

Generic Realism

The indifference of most political scientists to the validity of arguments, and the resulting incentive to develop competing brands of “theory,” leaves everyone with the impression that there is more to all the theories than is really there. Thus, although the brands are ostensibly in conflict with each other, they all actually give indirect support to each other. A recognition that they all shared the same law would mean that they would all have to go out of business, and better answers to the questions being debated might not lead to an easily marketable new brand. Instead, each is given credibility by its criticisms of the others, and a never-ending debate among the competing brands comes to define what the field is.⁷⁰ Moreover, ideas can be marketed under a brand name even if they have no logical relationship to each other, as offensive Realism, the security dilemma, and hegemonic stability theory all came to be thought of as part of structural Realism as defined by Waltz, simply because they all focused on the effects of anarchy.

Just as Chevrolet, Buick, and Cadillac are all General Motors brands, so offensive Realism, defensive Realism, hegemonic stability theory, and even structural Realism itself have all been considered brands of Realism, and they all gain some credibility from the common view that, if there is a law in this or that brand of Realism, we can nonetheless be pretty confident that, whatever the right answer is, it will turn out to be a Realist answer. But what is Realism?

It is sometimes said that prior to structural Realism there was classical Realism, which tried to explain interstate conflict as the result of an inborn human urge to seek power over other men—not surprisingly, this is sometimes called “human nature Realism.”⁷¹ However, this was a mistake, so the story goes, because violent conflicts are caused not by human nature but by anarchy, and therefore structural Realism corrects a mistake made by classical Realists. This is a story often used to explain the importance of Waltz’s writings, but it is a story that was first told by John Herz to justify the importance of his contribution to the subject.⁷²

However, “human nature Realism” seems on closer inspection to be a

70. The typical argument has the following form: (1) Either A or B is true; (2) B is not true; (3) Therefore A is true. The argument is valid, but the first premise is always false.

71. See, for example, Mearsheimer 2001, 18–19.

72. See Herz 1959, 232. In spite of this, both Keohane and Ashley have counted Herz among the classical Realists, and Herz’s security dilemma does not even appear in the index to

caricature. Its main academic exemplar is Hans Morgenthau, who wrote a famous textbook about international politics that is generally considered to be the main contribution to Realist doctrine prior to Waltz's *Theory of International Politics*. In the first edition of that book, Morgenthau wrote: "Domestic and international politics are but two different manifestations of the same phenomenon: the struggle for power" (1948, 21). But he also said that "the statement that A has or wants political power over B signifies always that A is able, or wants to be able, to control certain actions of B through influencing B's mind" (14). And that is why, "[w]hatsoever the ultimate aims of international politics, power is always the immediate aim": anything one wants to accomplish in politics of any sort requires getting other people to do something they would not otherwise want to do (13).

Thus Morgenthau did not say that people have an inborn urge to dominate other people (though admittedly he sometimes used language that could be interpreted in that way); he merely stated the truism that to achieve any political objective, no matter what it was, one had to be able to influence other people. The difference between international and domestic politics in this respect, he said, was that in domestic politics

Cultural uniformity, technological unification, external pressure, and, above all, a hierarchic political organization co-operate in making the national society an integrated whole set apart from other national societies. In consequence, the domestic political order is . . . more stable and to a lesser degree subject to violent change than is the international order. (Morgenthau 1948, 21)

It certainly sounds as though Morgenthau was talking not about the effects of human nature but about the effects of anarchy.⁷³

The other main "classical Realist" in the Realist canon is E. H. Carr, whose book *The Twenty Years' Crisis* (1946) seems to have given Realism

Keohane 1986, which is devoted to an evaluation of Neorealism (199, 257). This illustrates once again the confusion about what structural Realists think the explanation of war really is and what role the security dilemma is supposed to play in it.

73. For further evidence on this point, see Frei 2001, 140. There is a similar misinterpretation of Hobbes that is common. Hobbes said: "I put for a general inclination of all mankind, a perpetual and restless desire of power after power, that ceaseth only in death" (1957, 64). But he defined "power" as a man's "present means, to obtain some future apparent good" (56). And the reason men sought "power after power" was not "that a man hopes for a more intensive delight, than he has already attained to; or that he cannot be content with a moderate power: but because he cannot assure the power and means to live well . . . without the acquisition of more" (64). Hobbes was perhaps the first "offensive Realist." Morgenthau's emphasis on the struggle for power was, according to Frei, influenced primarily by the writings of Nietzsche and Weber, a fact that Morgenthau carefully concealed in order to avoid alienating his American audience (Frei 2001, chaps. 5 and 6).

its name. But Carr spelled the word with a lowercase *r*, and the message of Carr's book was that the thinking behind the League of Nations had been utopian and therefore unrealistic. This use of the term is consistent with one of the definitions of the word *realism* given by the *American Heritage Dictionary*, which is simply "The representation in art or literature of objects, actions, or social conditions as they actually are, without idealization or presentation in abstract form." John Herz agreed with this interpretation:

Strictly speaking, the terms "realism" and "idealism" should not be applied to theories. Theory is either correct or incorrect, depending on how it analyzes what happens in politics, but perhaps it is permissible to call a correct analysis a realistic one. Chiefly, however, the terms apply to actions and actors, those who behave according to "real," that is, existing givens, and those who engage in wishful thinking. (1981, 182)⁷⁴

In the first edition of his famous book, Morgenthau did not mention Realism or identify himself as a Realist.⁷⁵ And in *Man, the State and War* and *Theory of International Politics*, Kenneth Waltz did not identify himself as a Realist either—the terms *Neorealism* and *structural Realism* were bestowed on him by others. He did say that his ideas were "closely identified with the approach to politics suggested by the rubric, *Realpolitik*" (Waltz 1979, 117). But the word *Realpolitik* was introduced into the German language by August Ludwig von Rochau, a disillusioned participant in the failed revolutions of 1848, in a book about politics in the German-speaking states (Rochau 1859). The word meant, and still means, no more than the word *realism* means as described in the passage just quoted from Herz: Rochau came to believe that he and other revolutionaries had engaged in wishful thinking.

Thus, generic Realism seems to be nothing more than realism and therefore nothing very specific. If so, then the credibility of the Realist brand is entirely dependent on the inferences Realists have made about the effects of anarchy and the security dilemma, and if those cannot be justified then there is no more general Realist doctrine to fall back on. Any work that focused on war and military capabilities that reflected skepti-

74. However, see Herz 1951, where the terms *realism* and *idealism* are applied to theories.

75. He had, however, frequently used in his writings the English words *realism* or *realistic* (as well as their French equivalents) in criticizing writers who, he thought, had overestimated the efficacy of norms and the law in controlling social conflicts (Frei 2001).

cism about the prospect of eliminating war in the near future might call itself realist, but Realism itself would provide no justification for such skepticism.

That possibility seems consistent with the definition of the latest brand of Realism, neoclassical Realism, which was invented by Gideon Rose simply as a way of organizing a review of several otherwise unrelated books. "The works under review here," he says, belong to a school of thought

which I term "neoclassical realism." It explicitly incorporates both external and internal variables, updating and systematizing certain insights drawn from classical realist thought. Its adherents argue that the scope and ambition of a country's foreign policy is driven first and foremost by its place in the international system and specifically by its relative material power capabilities. This is why they are realist. They argue further, however, that the impact of such power capabilities on foreign policy is indirect and complex, because systemic pressures must be translated through intervening variables at the unit level. This is why they are neoclassical. (Rose 1998, 146)

The authors reviewed did not identify themselves as "neoclassical Realists," but given the importance of brand names in the field of international politics, the gift of a brand can hardly be refused, and thus it is not surprising that the term has had a life beyond Rose's review essay.⁷⁶ Instead of inventing new brands of Realism and new counter-Realisms, students of international politics should remove the word from their vocabulary, so that arguments can stand on their own and be evaluated independently of each other.

What Next?

An understanding of the deficiencies of structural Realism supplies us with a list of important questions that need answers. At the heart of all of them is the question of what explains the occurrence of war, whether interstate war or not. An answer to that question would help us understand the relation between government and war, which would in turn help us determine

76. See, for example, Schweller 2003. Note Rose's assumption that theories are defined by the "variables" that they focus on.

whether a world of sovereign states is doomed to war or not.⁷⁷

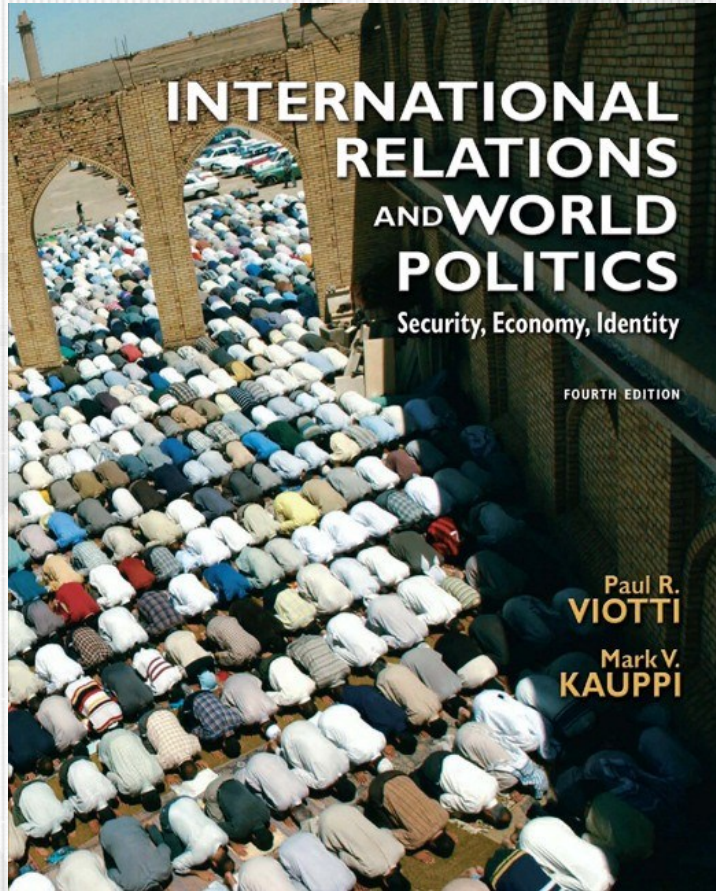
While Realism contains no valid arguments in support of answers to these questions, the works of Carr, Morgenthau, Herz, and Waltz are indeed part of a longer tradition of thinking about them. The proper name of that tradition, however, is not realpolitik but *Staatsräson*, *raison d'état*, *ragion di stato*, or reason of state. That is the tradition that Waltz claimed to be heir to when he invoked the name of Machiavelli in illustrating what he meant by realpolitik (Waltz 1979, 117). However, that is a tradition of thought not about international politics but about fundamental questions of political order that help us understand where the Western distinction between domestic and international politics came from. Supporters of the various “isms” have left quite a misleading impression of what writers in this tradition actually said. If we are to settle the issues that have been debated in the wars among the “isms,” we will have to take a closer look at this tradition. That will be the subject of the next chapter.⁷⁸

77. The failure of structural Realism to provide a clear answer to this question is the main theme of Van Evera 1999. However, there is actually no explanation of the occurrence of war in Van Evera's book, which instead contains a number of more or less plausible claims about what influences the probability that war will occur. This is like substituting a weather forecast for an explanation of the occurrence of snow. But if the weatherman does not know why snow occurs, one cannot have much confidence in his forecast. Van Evera claims that the hypotheses he discusses, which he calls “misperceptive (one-grained structural Realism,” provide support for the structural Realist “paradigm.” But it is actually not clear what they have to do with structural Realism. Many of the factors that Van Evera claims influence the probability of war will be discussed later in the book, beginning with chapter 4.

78. For background reading on this subject, I strongly recommend the following books, *to be read in this order*: Meinecke, *Machiavellism* (1998); Tuck, *The Rights of War and Peace* (1999); and Hirschman, *The Passions and the Interests* (1977). For a discussion of the influence of Meinecke on Hans Morgenthau, see Frei 2001.



Thinking Theoretically



What is Theory?

**Perspectives &
Approaches**

Major topics on Thinking Theoretically

- D Theory-building: what it takes to make the world more intelligible by explaining and anticipating or predicting the phenomena observed.
 - D Level of analysis: individuals and groups, state and society, and international systems or global society.
 - D Perspectives, standpoints, or images and approaches to understanding international relations or world politics.
 - D Interpretive understanding: challenges to long-established approaches to theorizing about international relations or world politics.
-

What is Theory?

- D A way of making the world or some part of it more intelligible or better understood, offering explanations or predictions.
 - D Normative Theory – the values or norms that inform judgments about what ought to be in international relations or world politics, as opposed to the way things are.
-



Levels of Analysis

D Individual level

- Human nature and psychology
- Leaders and beliefs systems
- Personality of leaders
- Cognition and perception or misperception

D Group level

- Government bureaucracies
 - Policy making groups
 - Interest groups
 - Other nongovernmental organizations
-

Levels of Analysis

D State and Societal (or National) level

- Governmental

- D Structure and nature of political systems
- D Policymaking process

- Societal

- D Structure of economic system
 - D Public opinion
 - D Nationalism and ethnicity
 - D Political culture
 - D Ideology
-

Levels of Analysis

D International – World (or Global) level

- Anarchic quality of the international or world politics
 - Number of major powers or poles
 - Distribution of power/capabilities among states
 - Level and diffusion of technology
 - Patterns of military alliances
 - Patterns of international trade and finance
 - International organizations and regimes
 - Transnational organizations and networks
 - Global norms and international law
-

Levels of Analysis viewed another way

Many influences affect the course of international relations. Levels of analysis provide a framework for categorizing these influences and thus for suggesting various explanations of international events. Examples include:

Global Level

North-South gap
World regions
European imperialism
UN

Religious fundamentalism
Terrorism
World environment
Technological change

Information revolution
Global telecommunications
Worldwide scientific and
business communities

Interstate Level

Power
Balance of power
Alliance formation
and dissolution

Wars
Treaties
Trade agreements
IGOs

Diplomacy
Summit meetings
Bargaining
Reciprocity

Domestic Level

Nationalism
Ethnic conflict
Type of government
Democracy

Dictatorship
Domestic coalitions
Political parties and elections
Public opinion

Gender
Economic sectors and industries
Military-industrial complex
Foreign policy bureaucracies

Individual Level

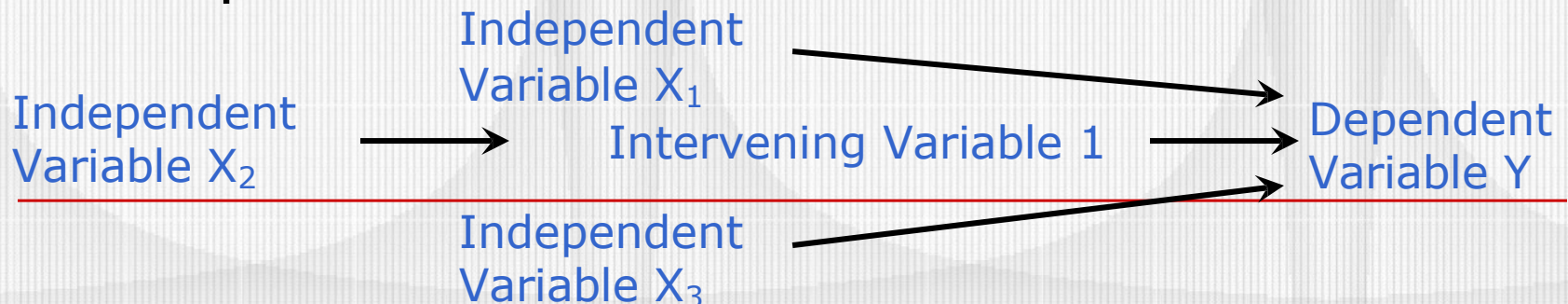
Great leaders
Crazy leaders
Decision making in crises

Psychology of perception and decision
Learning
Assassinations, accidents of history

Citizens' participation (voting,
rebellious, going to war, etc.)

Using Levels of Analysis to Formulate and Test Hypotheses

- D** Dependent variables – the unit of analysis (what is trying to be explained)
- D** Independent variables – factors at different levels of analysis casually related to what is trying to be explained
- D** Intervening variables – other factors which may facilitate or block the effects of independent variables or each other or on dependent variables





Perspectives & Approaches

1. Realism focus on power and balance of power among states in world politics.
 2. Liberalism (pluralist perspective) looks at not only states but the roles individuals, international and nongovernmental organizations play.
 3. Economic structuralism emphasizes the importance and roles of material factors or structures as capital-owning classes or capital-rich countries in international relations.
 4. English School “rationalist” portray international or global society using realist ideas emphasized in writings by Machiavelli, Hobbes, or rules and institutions in writings of Grotius.
-



Alternative Perspectives & Approaches

1. Constructivism underscores the importance of ideas and in establishing or constituting the world we understand.
 2. Feminism focuses on gender to understand not only politics within state and society, but also international relations or world politics.
 3. Postmodernism emphasizes the subjective dimensions of what is found in the world around us. This view deconstructs or unpacks spoken or written communication to find underlying meanings.
 4. Critical theory probes beneath the cover stories given to reveal underlying calculations of interest or exploitation.
-



Realists Perspectives & Approaches

1. As the term suggests, this outlook views the world as it is, not as it ought to be.
 2. Stability is key in political power, gaining and staying in power.
 3. With order, other objectives are possible, i.e. economic and social.
 4. In IR, realists emphasize the struggle for power and influence among states and empires.
 5. Security may extend to regional or world conquest, a dynamic encouraging competition and violence.
-

Realists Perspectives & Approaches

6. Competition and warlike conditions in IR result from combined factors:
 1. Human nature or individual psychology and small group dynamics
 2. The nature of states or societies
 3. The structure or distribution of power among states

Structural realists or Neorealists – look at the distribution of power among states of structures as affecting behaviors within an international system (unipolar, bipolar, multipolar).



Realists Perspectives & Approaches

- D Power is important in achieving objectives and interests of states.
 - D International organizations, alliances, and other coalitions are merely mechanisms used by states to exercise power.
 - D For some states the pursuit of power is not just a means; it becomes an end in itself.
 - D Great powers actively construct or change power balances to assure their best position.
-



Liberals - Pluralist Perspectives & Approaches

1. The term Liberal is more commonly used and is not to be confused with the political ideology found on the left of the U.S. political spectrum.
 2. Both share the idea that a complex global society comprises of multiple kinds of actors.
 3. Liberalism is a more inclusive outlook, with states as key players but not the only ones in world politics.
 4. International organizations, i.e. United Nations and European Union, are both arenas where states compete for influence and independent actors seeking to shape the global issue agenda.
-

Liberals - Pluralist Perspectives & Approaches

- 5. View states as battlegrounds for conflicting bureaucratic interests subject to domestic and transnational interest groups.

Neoliberal institutionalists

- D View the state as an enduring actor.
 - D Non state actors (international & no governmental organizations) are important too.
-

Economic Structuralists Perspectives & Approaches

1. both Marxists and non Marxists
 2. tend to focus on relations of dominance in society in the form of economic structures (national or global) that purportedly also contribute to explaining the conduct of both state and non state actors.
 3. divide world between core – periphery, north – south, capital rich v. capital poor countries
-

English School “Rationalists” Perspectives & Approaches

1. This school of thought draws on elements of the other perspectives, specifically realism and liberalism.
 2. View international politics in terms of interactions within a worldwide or global society.
 3. Emphasize global civil society, international law, moral and ethical principles.
 4. Associated with the writings on Hobbes, Grotius, Kant, Machiavelli.
-

Constructivists & Feminists

1. Constructivist or social constructivism argue
 1. that states don't simply react to their environment, but dynamically engage it;
 2. that anarchy in international politics is what make of it.
 2. Feminism
 1. may share an affinity with social constructivism
 2. involves using gender as a category of analysis, focusing on differences between feminine and masculine understandings and actions;
 3. holds that gender permeates social life, has a largely unnoticed effect on the actions of states, and includes conceptions of masculinity.
-

Postmodernists & Critical Theorists

1. Critical Theorists & Postmodernists assume that facts, concepts, and theories may not be separated from values, as all 3 stem from their observation and construction by human agency.
 2. Both may be viewed separately because most critical theorists retain strict methodological criteria to guide their work.
-

Glossary List

- D Positivism
- D Modernism
- D Third World
- D Compadre class
- D epistemology
- D ontology

The chapter concludes with a brief overview of normative (value) considerations as they apply to world politics. Much of the theoretical material in terms of specific examples is consigned to boxes.



Review – How much do you understand?

1. A focus on mechanisms of domination in a capitalist world order is most closely associated with

- A. economic structuralists.
 - B. liberals.
 - C. realists.
 - D. social constructivism.
-



Review – How much do you understand?

2. Political scientists who view the world in terms of a multiplicity of different kinds of actors engaged in transactions are called

A. feminists.

B. liberals.

C. realists.

~~D. economic structuralists.~~



Review – How much do you understand?

3. What are the four theoretical perspectives used to explain why the U.S. invaded Iraq?

- A.** Individual level (psychological factors) – individual roles and personalities of president, VP, cabinet secretaries sharing ideological worldview or neoconservative perspective.
 - B.** Group level (collective role of policymakers) – White House, Defense Secretary, and trusted advisers favored preemptive military intervention.
 - C.** International, World (or Global) level – unipolar international system with the U.S. as sole superpower facing no obstacles.
 - D.** All of these answers
-



Review – How much do you understand?

4. Global civil society refers to

- A.** rule of international law.
 - B.** voluntary institutions.
 - C.** networks.
 - D.** All of these answers
-



Review – How much do you understand?

5. Those in the English School tend to see international politics in terms of

- A.** class struggle conducted on a global scale.
 - B.** interactions within a worldwide or global society.
 - C.** global ideals that ultimately will displace realist balance-of-power understandings.
 - D.** An almost mechanical system of states.
-

P&IR Dissertations



2013-14

Writing your Politics & IR Dissertation

DR SARA DORMAN
&
DR MICHAL ROZYNEK

IN WRITING YOUR DISSERTATION, YOU WILL DRAW UPON ALL THE SKILLS YOU HAVE BEEN DEVELOPING SINCE YOU STARTED STUDYING POLITICS & IR – READING CRITICALLY, ANALYSING ARGUMENTS, COLLECTING DATA, AND WRITING EFFECTIVELY.

THE DISSERTATION REQUIRES YOU TO DEVELOP YOUR OWN RESEARCH QUESTION, FOCUSING ON A TOPIC OF INTEREST TO YOU.

YOU WILL NEED TO SET YOUR OWN DEADLINES FOR WRITING DRAFTS AND MEETING WITH YOUR SUPERVISOR.

SEE YOUR SUPERVISOR AS A RESOURCE: THEIR JOB IS NOT TO TELL YOU WHAT TO DO, OR TO SAY ‘THAT’S GOOD ENOUGH’ BUT TO HELP YOU MAKE YOUR WORK BETTER – BY HELPING YOU FOCUS YOUR QUESTION, LOCATE THE BEST SOURCES, AND DRAFT MORE EFFECTIVE CHAPTERS. MAKE SURE YOU SCHEDULE APPOINTMENTS OR ATTEND OFFICE HOURS.

THE POINT OF SENDING A SUPERVISOR A DRAFT CHAPTER IS FOR THEM TO SUGGEST WAYS FOR YOU TO IMPROVE IT. THERE IS NO POINT IN SUBMITTING A DRAFT TO THEM, UNLESS YOU MAKE TIME TO CONSIDER THEIR SUGGESTED CHANGES.

WRITING A DISSERTATION IS A BIG CHALLENGE FOR MANY STUDENTS, BUT IT CAN ALSO BE VERY REWARDING, IF YOU PUT IN THE TIME AND EFFORT!

Writing your Politics & IR Dissertation

2013-14

Full-class meetings

You are expected to attend the full-class meetings convened by the Dissertation Conveners. These are intended to convey important information necessary to the smooth development and completion of your dissertation, as well as offering the opportunity for an exchange of ideas between the students as a group, and with the Dissertations Conveners.

Small group workshops

There will also be small group workshops organized three times through the year where you can work on practical skills with other students and tutors. Likely topics will include writing literature reviews, structuring your dissertation, effective writing and writing an abstract. The times/places for these will be announced on the Learn page.

The Dissertations Conveners:

The dissertations conveners are responsible for allocating supervisors, providing general guidance and advice, running dissertation meetings, and resolving any difficulties that may arise.

Supervisors:

Students are supported by a member of staff who acts as supervisor. Supervisors may be able to give advice on practical issues such as: the subject and title of the dissertation, its organisation and structure, and on source material and a bibliography.

Supervisors can also be expected to comment upon dissertation outlines, chapter plans and timetables, and to provide feedback on around two draft chapters – about half the dissertation- in good time -- normally within two weeks of receipt.

Supervisors are not expected to direct your work or to comment on the final draft: a dissertation is intended to demonstrate a student's ability to work independently. Supervisors have other commitments, and time must be allowed for them to read and provide feedback on your work.

Students' responsibilities:

You need to find out when your supervisor is available and how soon before a meeting they require written submissions (if they don't specify this, assume 5 working days) and how soon they expect to be able to respond, as a rule, to you. It is reasonable to expect supervisors to respond to email queries within 5 working days (unless they are away), and provide detailed comments on written work within two weeks of receiving it. This means that you should not expect to be able to submit chapters to supervisors less than 2 weeks before the final submission date.

Staff will not normally be available to provide supervision or written feedback during the Christmas and

Easter Vacations. It is up to you to make arrangements to meet your supervisor - they will not chase you for progress reports or draft chapters. Students should keep a log of meetings with supervisors, including dates of meetings and summaries of discussions.

If you have any difficulties that cannot be resolved between you and your supervisor you should contact the Dissertation Conveners.

Announcements

Course announcements will primarily be made via the Learn page and by email. It is your responsibility to keep track of these.

Draft Chapter Submission

Students are required to submit at least one draft chapter to their supervisor by the end of the Autumn Semester Monday 2 December 2013. Failure to submit this draft will incur a 10 mark penalty on your overall dissertation grade.

The draft chapter will not be given a numeric mark, it will be assessed as 'pass/fail' and you will be given feedback on it.

When sending a chapter to your supervisor, please name your chapter files something like 'surname_draft_chapter_5.doc' rather than 'my-chapter.doc' or 'thesis_draft.doc'

Marking

Dissertations are marked by two members of the Politics & IR subject area, one of whom is usually your supervisor. Your dissertation may also be read by the external examiner. Dissertation grades are provisional until the Exam Board has met in June.

SUBMITTING YOUR DISSERTATION

The dissertation deadline is 12 noon on 7 April 2014.

Dissertations will be submitted on-line only in 2013-14. Further information will be provided in Semester 2.

Please see the 'Honours Handbook' for further information on submission of coursework; Late Penalty Waivers; plagiarism; learning disabilities, special circumstances; common marking descriptors, re-marking procedures and appeals.

TECHNICAL SPECIFICATIONS

Font: For legibility, we prefer you to use a *Times* or similar font in 12 point.

Spacing: You should use at least one-and-a-half line spacing, and leave ample margins.

Page numbers: Pages should be numbered.

Length: The maximum length of a dissertation is **10,000 words**, including footnotes but excluding the cover page, acknowledgements, acronym/abbreviation list, bibliography, Abstract, ToC, and Appendices.

Electronic copies of dissertation can and will be checked to confirm the declared word lengths. 2 marks will be deducted for every 100 words or part thereof over the word limit. (There is no '10%+/- allowance/grace).

PLAGIARISM:

Please read the Honours Handbook regarding matters of **plagiarism**. Dissertations must be your own work. Your dissertation will be assessed for plagiarism using state-of-the art detection software, and penalties will be applied if necessary. Do not take any risks. Use a consistent system of referencing and acknowledge all sources

Dissertation Assessment Criteria

The following are the criteria through which the Dissertation will be marked. However, it is important to note that the overall mark is a result of a holistic assessment of the assignment as a whole.

Does the dissertation have a problematique, research question or hypothesis which it attempts to answer?

Does the dissertation effectively engage and show knowledge of relevant academic literature?

Does the dissertation demonstrate a logical and effective pattern of argument?

Does the dissertation effectively use primary or secondary evidence in support its argument?

Does the dissertation demonstrate reflexivity and critical thinking in relation to arguments and evidence?

Is the structure of the dissertation well-chosen and effectively implemented?

Is the dissertation adequately presented in terms of: correct referencing and quoting; spelling, grammar and style; layout and visual presentation?

GETTING STARTED

Most of you will already have identified a 'topic', but there is a lot more to a dissertation than that. You need to find a research question, identify relevant literature, collect data and plan a schedule.

What is a dissertation?

A dissertation is a 10,000 word written paper, which involves an in-depth exploration of a particular topic. Research can be wholly or largely library-based (involving secondary literature) or can involve empirical data collection (for example through surveys or interviews).

A dissertation is worth 40 credits and *is equivalent to two semester-length courses*. You should therefore be putting as much time into it as you do to any other year long course.

Ask yourself this: have you put as much time and effort into your dissertation as you have into taught classes where you need to turn up and participate in seminars or tutorials on a regular basis?

On any topic, there are huge numbers of books you could be reading. Remember that you don't have to read everything. You need to be self-disciplined and know when to stop. You are aiming to answer your research question, not discover all there is to know about your broad research subject.

In guiding your time and task management, it's useful to have a chapter outline. This needn't be something you stick to rigidly - you might want to play around with the structure at the writing-up stage if you feel this would make it stronger. This is much easier to do once the work has been done! What a chapter outline will do at this stage is ensure that you are covering everything that you need to cover. This should really be an extension of the proposal you discussed with your supervisor.

Draw up a list of tasks that you have to do, and draw up a realistic timetable in which to fulfill these tasks. (Examples of a task: to examine government policy on social security, etc., to compile election results, etc.). Then give yourself a realistic timeframe in which to complete this task.

Draft a timetable according to the chapter outline. But be realistic. There is no point in setting goals that you will never attain. There is no point, for example, in saying you'll have written two draft chapters by the beginning of next term if you know that you're going to be spending Christmas and New Year partying with friends or visiting family! Setting goals that you never attain can also be very demoralising.

At the same time, you should build in time for slippage. It may take longer to do a particular piece of work than you planned. You also need to make sure that you leave enough time for writing-up at the end.

Remember that you are not the only student being supervised by your supervisor, and that she or he will have many other commitments as well. You can't expect to hand in a piece of work and get it back the next day. So you should ensure that you leave yourself enough time to submit work and get it back from your supervisor, and be able to consider the suggestions she or he has made.

You should leave yourself at least two weeks at the end to edit your final draft. Editing is as important as writing.

There is no right way to go about this, but our advice would be to write as you go along, rather than to leave it all until the end. Writing helps you think, as well as the other way around. By writing, you formulate your thoughts and refine your ideas. This should all be part of the research process. You should aim to have at least one draft chapter written by the end of the Autumn term and a second chapter written by midway through the Spring Term. Your supervisor will provide you with detailed feedback on these chapters.

Who should write a Politics & IR dissertation?

Dissertations are compulsory for students taking single Honours Politics or single Honours International Relations and are open to all those taking Politics as part of a Joint Honours degree.

A number of Joint Honours degrees have a compulsory dissertation project, which students may opt to write either in Politics or in their other subject area. You must decide as soon as possible and make sure that you are registered for the correct course by September 2013 at the latest.

NB: Some joint honours degrees require a Politics dissertation (eg Modern Languages); some do not allow students to write Politics degrees (eg Arabic)

Please check Degree Requirements in the Options Booklet, the online DRPS or with an SSO or your Personal Tutor to confirm which regulations apply to you.

Choosing a topic and writing a proposal

A brief research proposal must be submitted online via a link which will be emailed to you. This will enter your proposal onto a secure database.

This should specify the topic you wish to work on, and a preliminary research question. You will also be asked to identify any courses which you have taken which relate to the topic.

You will find out who will supervise your work in September via the Learn page for your dissertation course.

It is your responsibility to make contact with your supervisor and arrange your first meeting. You can email them and/or attend their office hours.

When you email them, it is helpful to introduce yourself and explain that you are writing an Honours dissertation. You may also want to send them a short – 1-2 pages – outline of your topic and proposal.

Refining your topic and preparing an outline

You may find it helpful to prepare a **1000-word** outline of your proposed topic for submission to your supervisor at the beginning of the Autumn Semester. This should develop on the topic you submitted in week 1. It should include:

Proposed title of dissertation

Topic area (e.g. US politics, gender politics, human rights etc)

Research question (or questions)

An initial chapter outline

The sources to be consulted – are you using primary data? Do you need to complete an ethics self-assessment?

A preliminary discussion of relevant literature – this may be a simple bibliography, an annotated bibliography, or a few paragraphs describing the literature.

A work plan setting out your timetable

Useful resources

We recommend that you consult Karen Smith et al, *Doing your Undergraduate Social Sciences Dissertation*, which is available via the library as an e-book and in hard copy.

A very good on-line resource about researching and writing undergraduate dissertations in the social sciences is also found at: <http://www.socscidiss.bham.ac.uk/>

Other similar guides include:

Harrison, L (2001), *Political research: an introduction* (London: Routledge)

Murray R (2002), *How to write a thesis* (Buckingham: Open University Press)

Oliver P (2004), *Writing your thesis* (London: Sage)

Silbergh, D (2001), *Doing Dissertations in Politics: a student guide* (London: Routledge).

SOURCES, DATA AND QUESTIONS

Your dissertation depends on an argument which answers your research question, and links together the various chapters. In order to develop and defend this argument you need data or evidence, which supports your analysis. This section will help you find sources, select data and develop your argument.

Finding sources

Coursebooks: Look at the reading lists for courses you have already done that relate to your topic. They should guide you to materials that your instructor thinks are good quality.

Supervisors: Ask your supervisor to recommend readings. Depending on your topic, they may or may not be able to do this.

Shelfmarks and Search terms: if you find one book that you already know is relevant, this will lead you to the relevant shelfmark where you will find books on a related topic. The word cloud on searcher will also help you find related search terms.

Bibliographies: Books and articles that you have read for coursework will make reference to other books that relate to the topic. Use the references in endnotes or bibliographies to lead you to relevant works.

Journals: You are expected to use journal articles as well as books in your literature review. Identify the relevant journals in the library that relate to your topic. Examples of useful journals in politics are: Nations and Nationalism, Regional and Federal Studies, Scottish Affairs, Elections, Public Opinion and Parties Yearbook, Parliamentary Affairs, Political Quarterly, International Affairs, West European Politics, etc. Often in bound volumes, one issue will have the index for all of the issues of that year, or they'll all be at the beginning. Sometimes, the index will be found in a separate slim volume.

Remember that JSTOR is just an archive, like the library shelves. It doesn't include the most recent issues of journals. You need to search databases as well as JSTOR to find recent articles.

Databases and gateways:

Use 'searcher' via the Library homepage to search lots of databases at once, or you can browse the alphabetical listing under 'databases'. Some useful databases include:

Firstsearch/ Articlefirst. – an extremely useful database. You can search under subject or keyword. It will show you many options. Some will be available on-line, some will be in the Edinburgh University, others will not. Another Firstsearch database is Worldcat, which lists most books published, you can search it in the same way, and using the same searches as Articlefirst.

ASSIA: An international abstracting and indexing tool for health and social science professionals. It provides abstracts from around 650 UK, US and international journals.

Web of Science (<http://wok.mimas.ac.uk/>) The social sciences citation index is likely to be most useful for politics dissertations. You can search on-line for articles published in thousands of journals since 1956. Like Articlefirst, it will give you thousands of returns, you may want to narrow down your search to just focus on the past 5 years, or similar.

Googlescholar: will help you find recent journal articles but it is not as comprehensive as a database like Articlefirst or WoS. Nor does it link you to our catalogue.

Social science gateways are often a very useful way to access relevant material. See, for example, <http://www.jisc.ac.uk/subject/socsci>

Inter-library loan: if a book or article is not available in either the University Library or the National Library of Scotland, you can ask your supervisor to write a letter of support to take to the library. This should entitle you to request an inter-library loan. Each request will cost you £4, which is subsidized by the University. More information is available on <http://www.lib.ed.ac.uk/services/ill1.shtml>. It is helpful if you know that your book or article is available from a UK university. Check COPAC <http://copac.ac.uk/> for the major UK universities or the British Library catalogue. Photocopies of articles or books held at the BL usually take 1-2 weeks to arrive at most, often they are scanned and emailed within a few hours. ILLs can also come from abroad, but this may take longer.

Internet – many organisations in the field of politics now have their own websites, and publish documentation on their web pages that is freely accessible and can be downloaded, but you should be aware that such data is not necessarily reliable.

Masters and PhD dissertations: Can be searched on-line through databases available on the library website. They can be ordered as ILLs.

For students writing a dissertation on development-related topics

The Governance resource centre at Birmingham, <http://www.grc-exchange.org/> you can search the site, or look at particular themes. They have also compiled very useful subject lists of materials on various themes from public expenditure management to water and sanitation. The British Library for Development Studies at IDS also has a searchable on-line index – accessible from Searcher -- and many on-line documents at <http://blids.ids.ac.uk/blids/>. Both these services will send documents to you for a fee– but check, you may find that the library has a subscription, or that it is cheaper to get on ILL. So, for example if you search under NGOs, you would find an abstract of [Moore, M. and Stewart, S. 2000, 'Corporate Governance for NGOs?,' in D. Eade, ed. Development, NGOs, and Civil Society, Oxfam, Oxford, pp. 80-90.](#) which you can also find in our library.

Collecting and Using Data

If you intend to collect and use primary data, make sure you discuss this with your supervisor before you start your research. Some of our best dissertations have been based on primary research, but you will need to plan ahead. Primary data may include interviews, statistical data from surveys, analysis of newspaper articles, or your observations of an organization or event (eg observing participants on a protest march).

If you need special skills or training (eg in analyzing quantitative data), you should consider taking some research methods courses (see the honours options handbook). If you are going to conduct interviews or focus groups, or carry-out participant observation, you need to contact people in advance and allow sufficient time. You will also need to carry out a research ethics self-evaluation – your supervisor will help you with this.

If you intend to base your dissertation on a gap year or a summer placement, talk to your supervisor about this before you leave. And read the section on research ethics below.

Research Ethics

All research projects must be assessed for their ethical implications. This includes undergraduate dissertations if you are collecting primary data. If you suspect that this might apply to you, please discuss this with your supervisor, who will assist you in determining what needs to be done to ensure that your research is ethical.

Details of the process can be found on the School webpage:

http://www.sps.ed.ac.uk/admin/info_research/ethics

Reading Academic literature

When you read academic literature, you are not just mining it for ‘facts’ but also exploring arguments and debates. So, be sure to think about the context – when was the book or article written? by whom? what debates was it engaging with? How did its methodology or theoretical approach shape its contribution?

You may find it helpful to look at book reviews, or follow up articles that cite the item you are reading – Google scholar is very useful for doing that.

It can also be helpful to prepare an annotated bibliography – which just means making notes on what you’ve read – to help you clarify your thoughts before you write the literature review (which may or may not be the first chapter you write)

Reviewing Literature

Reviewing literature is important because it helps you narrow down your topic and formulate your research question. It may become a chapter or a part of a chapter of your dissertation. **Not every dissertation will have a chapter called ‘the literature review’ but every dissertation must show familiarity with the relevant literature.**

Before embarking on the literature review, it is imperative that you have clearly identified your topic, and that you have ‘unpacked’ the topic into its component parts. This should have been the task of the **Dissertation Outline, which you have discussed with your Supervisor in Week 3**. A good and focused outline will ensure you are aware of the nature of the problem to be examined, the main research questions stemming from this, the major concepts you want to examine, the theoretical issues raised by your topic, and the empirical questions you need to address in the process of your research.

Whether your dissertation is theoretical or empirical – you still need to have – and demonstrate you have – knowledge of existing literature. Literature review is an ongoing process – you will go back to it when trying to address research questions and understand the data or findings, or elucidate your thoughts. But it is usually expected that you spend a fair amount of time at the beginning familiarising yourself with the literature relevant to your field.

The literature review serves a number of inter-related functions:

- It helps you refine the problem set out in the outline, and to build upon it
- It allows you to develop an in-depth understanding of the subject area in which you are working, and to learn from existing research in this field. This in turn should give you food for thought that informs your own thinking and your own research.
- It provides guidance in addressing some of the research questions set out in your outline – it won't always be necessary to conduct empirical research from scratch if previous research has answered some of these questions for you. Remember there is a limit to what you can do by way of empirical research anyway (time, resources, expertise).
- It helps you gain an understanding about the theoretical and analytical debates that are prominent within current work on your topic. You are expected to **engage** in these debates.
- It will help you make sense of your data and findings when you gather them.

As well as summarising the views and conclusions expressed by established academics, you are expected to be analytical, to weigh up different arguments and points of view, and to critique existing work, where you take issue with it, but avoid saying 'I think' or 'my opinion is' – it is better to express your opinion in terms of your argument, using evidence to defend your position robustly.

This adds scholarly weight to your dissertation. We are looking to see that you have a good understanding of existing literature, that you can see where your work fits in the academic field. We want not only to see that you've read books, articles and reports, but that you've thought about them, given consideration to them, addressed the merit of existing work, maybe even challenged them as a result of your own research.

Three years of University essay writing has already given you most of the skills you need to undertake this task. The difference now is that whereas in essay writing you are usually addressing a set question and using a set reading list in so doing, with the dissertation you set your own question and build your own reading list. There are a number of cautionary points that you should take note of:

Always retain focus – this is why it is important to have a good research outline that you have discussed with your supervisor. There is no point in aimlessly reading and reading around the general area and hoping this will bring focus to your research.

Don't waste time on irrelevant material – if half way through a book, you aren't getting much use out of it, ditch it! Don't feel you have to get to the end of a book if it isn't really relevant. Likewise, resist the temptation to use a book just because you read it. We are always looking to see that you have demonstrated the relevance of the material you have read, i.e. how it relates to the specific topic you are addressing.

Use academic sources and specialist texts. Textbooks may be useful to direct you to the specialist texts, and familiarise you with the key debates, but in themselves, they are not very useful sources – they are too general for this level. The Internet is useful, especially if examining an ongoing issue where the academic textbooks may

be behind events, but the internet is never going to be good enough by itself. You **must** embed your research and your written work in the context of the existing academic literature and debates.

Don't try to read everything! Part of the challenge of the literature review is selecting the relevant material. You are not expected to have read every book or journal article that relates to your topic. You are expected to have read the key ones. Supervisors can help you identify the key texts. Occasionally, you find Readers on the areas that you are working on (for example, Readers on nationalism or the EU). They tend to include chapters from the key contributors to the field.

Identifying Your Question/Argument

Dissertations need to be coherent. The central research question or argument is what links together your various chapters. The different sections and chapters of the dissertation will all relate to one another, because they help you explore and explain a specific research problem. So it is vital that you identify and then focus on the central question at the heart of your dissertation, and the related sub-questions that derive from it: what's your point? What's the point of your dissertation?

At this stage, it is important that you are clear about the specific nature of the problem you are going to research. The literature review should help you do this. Having a clear research question, or set of inter-related questions, helps to guide your research and reading. In other words, what questions will you have in your head when conducting your research? To what end is your research serving? What are you looking for?

The 'what's your point?' question is a really useful question to ask yourself throughout the process of research and writing your dissertation. This can also be asked as 'So?' or 'So what?' Such a question helps to determine the relevance of the research, the relevance of the information being gathered, and the relevance of the paragraphs you will eventually be writing. If you ask this, and see no relevance, then forget it because it probably isn't relevant!

Define the problem. What is the central hypothesis or question under consideration? This is a key contrast with essays, where the question is set for you. Here you have had to do it yourself, within a broad area of interest (gender and media, Hutton Inquiry, etc). So, you should provide a **statement of the problem** – this may include a hypothesis (that A causes B) to be tested, and a central question to be addressed. NB this should be framed in such a way as to allow you to be analytical. Don't just ask 'how does the electoral system of country X work?' or 'what types of humanitarian intervention did country X engage in, or country Y receive'. Questions like that would lead you to a more descriptive account of a policy, or system. This question wouldn't answer the 'what's your point' or 'so what' question.

What is purpose/aim of the research? Is it to examine a policy, evaluate or develop a theory, analyse the significance or consequences of a political development/event, or what?

Why have you chosen this problem for consideration? What has prompted the research? (A change in legislation? A political development? A gap in the literature you have identified?). What is it that makes this an interesting area of study, and an important question to be examined? Try to be precise. It's not enough to say you were curious about how parties fight elections or you were wondering about the effect of devolution. There is no research problem in wondering how something happens or develops. Your statement of the problem must be precise and one which will be examined in the course of your research.

What theoretical approaches help you understand or explain your material? How does it relate to bigger theoretical arguments within political science? Is there an over-arching theoretical framework against which your work can be situated?

What are the sub-questions – the research questions – this has generated? A hypothesis or statement of the problem will have generated a series of research questions to test your hypothesis or examine the problem. Write down the 2 or 3 that are guiding the research that you will be embarking on this term. Remember here that you have to ensure that you have considered alternatives to your explanation. You may also need to look at the controlling variables, e.g. for a topic on women in parliament it is not enough to look only at the women, but you would also need to look at men.

WRITING, ANALYSING AND REFERENCING

Finally, you need to pull together your argument and evidence into a bundle of chapters which will make up your dissertation!

Writing, reading and revising all take place throughout the year, but at some point you will 'start' to write your first chapter. Many people start with a **literature review**. This helps you identify the relevant literature and identify your research question. Other people find it more helpful to start writing a **case study chapter**, and then write the literature review when they are more confident about the topic. Some dissertations, for example in Political Theory, integrate the literature review into all chapters, even if there is no chapter called 'THE literature review'.

You should discuss and agree with your supervisor which is best for you. In the end, your dissertation must show that you are familiar with the literature, whether or not you start by writing a literature review or not. The important thing is to start writing. Remember, you must submit a draft chapter in Semester 1.

STRUCTURE, ANALYSIS AND ARGUMENT

We have stressed the need to be clear about the purpose and point of your dissertation. It is important when you go into the writing stage that you have a clear idea of what the dissertation is about. What is the purpose of the dissertation, what is the problem under examination, what is the central contention or argument you are developing, and what central questions does your dissertation raise and address? Having a clear idea of precisely what it is you're examining is essential if you are to build a sound structure.

Set aside sufficient time for the data analysis process and for drafting and redrafting chapters. Keep questioning yourself and your material to identify how it addresses your research questions, and how the material inter-relates. Be prepared to make critical and balanced judgment of the material. Treat data critically and remain focused. Avoid going off at tangents that detract from the purpose of your study. Avoid making exaggerated claims to 'proof' or drawing conclusions, which your material does not support.

You might find it useful to adopt a system of coding to categorise your material into chapters and sub-themes within chapters, to facilitate the analytic process. In doing this, ensure that your analytical categories are consistent with the aims and objectives of the dissertation.

The structure of the thesis is about **giving shape or form** to the mass of data and material you have gathered in your research and begun analysing in the data analysis phase. In structuring the dissertation, you need to select the data that is relevant and put it together in a well-organised and coherent manner. There is no set structure that should be adopted. However, the structure you choose must allow you to examine the different aspects of the problem, idea or issue under study, and each chapter must have its place in relation to the overall theme, and together the chapters must hang together to form a coherent structure and argument, where the theme has been developed throughout, leading us to the conclusion. Clearly then, deciding how to put the

material together and determining which structure to adopt can only be decided once you have a clear idea of the purpose of the project.

Selecting only relevant material: Students will often feel a very natural and human desire to demonstrate the full extent of their labours, to prove how much they've read and how much work they've done. One of key skills in dissertation writing is in selecting only the material, which is relevant. That means being prepared to leave out material which doesn't fit – which may be superfluous or unacceptable in some way. So, think of it as a journey - avoid temptation to meander down wee side roads that may be interesting but take you off the main track. Leave out what's irrelevant to the central problem being addressed.

Introduction - all dissertations must have an introduction. This should include: A statement of the problem or main theme of the dissertation. A clear expression of the purpose of the dissertation and its rationale. The research questions it has generated. The methodologies or approaches used in examining these questions. Why this is an interesting topic of study and what it adds to scholarly work. Signposts - how the direction of your argument and how the rest of your dissertation will develop.

Conclusion - all dissertations must have a conclusion. This should be used to sum up the main points of the argument and pull the threads together. It returns to statement of problem and purpose of the dissertation and draws conclusions in light of evidence presented. Your Conclusion must be consistent with the rest of your argument. It should not come as a surprise to the reader.

Middle Bit – there are no set rules about how to structure your dissertation. But the outline and purpose of the dissertation may suggest a particular structure.

The key points with respect to structure are that each chapter must have its place. The relevance of each chapter to the main problem or theme must be clearly explained. Chapters must hang together and flow one into the other. The reader must be able to follow in a logical sequence.

EDITING AND PROOF-READING

A dissertation is different from an essay because you receive feed-back on your chapters that will help you revise them. You should also read your own drafts critically and think about how you can make them clearer, more effective, and better.

Editing is as important as writing – and there is a lot of scope using 'cut and paste' to move things around, make amendments to structure, etc. Leave time to get it right!

Leave yourself a few days for a break. Put your draft chapters away in a drawer and then come back and read them as though they were someone else's. It is very difficult to criticise your own work (people can be too harsh and too kind to themselves). Try to be an impartial critic. Make an arrangement with fellow students to comment on each other's drafts.

Read your work aloud - this helps expose awkward sentence structure and unclear expression, and helps guard against repetition, poor construction of argument, etc.

Students often say 'I can't cut it down to 10 000 words'. Remember, you are being assessed on your ability to write concisely and effectively on your topic. Dissertations should not include extraneous text, unnecessarily wordy sentences, or show evidence of inadequate editing.

You are being assessed on your ability to write concisely and effectively on your topic.

REFERENCING

Finally, it is essential that you ensure that your work is properly referenced. This enables your readers to see that you have consulted and engaged with secondary literature, and also to identify the sources of your primary data. It is also an important component of academic honesty that helps ensure you are not vulnerable to accusations of plagiarism.

It is important that you choose a good system of referencing and use it consistently. Ensure that you begin keeping careful records of your sources right from the start of your project.

Detailed advice on referencing is available on the Learn page, the recommended textbook and many other reference manuals available via the library.

You may find a system such as Endnote or Zotero useful. Please ask the dissertation conveners if you would like further advice, or if a small group workshop on referencing would be helpful to you.

SCHEDULE: IMPORTANT DATES AND DEADLINES FOR YOUR DIARY

March 2013	Full-class dissertation preparation meeting
May-June	Preliminary small group meetings: finding a research question
September	Students submit preliminary topics and questions (via Learn). Supervisors allocated
18 September	Full class meeting LT 1 Appleton Tower
September/October	Meet with supervisors
24-25 October	Small group workshops Appleton Tower M2a/M2b
October-November	Students write proposal, timetable, or lit review, or annotated Bibliography, & start on draft chapter.
27 November	Full class meeting LT 1 Appleton Tower
2 December	<u>Submit a draft chapter</u> c. 1500 words. Pass/Fail assessment. 10 mark penalty for failure/non-submission.
Jan-Feb-March 2014	Meet with supervisors to discuss progress, submit 2 nd draft chapter
23/24 January	Small group workshops Appleton Tower M2a/M2b
5 March	Full class meeting LT 1 Appleton Tower
20/21 March	Small group workshops Appleton Tower M2a/M2b
	<i>Please remember, you should not be sending draft chapters to your supervisors in the last 2 weeks before the deadline.</i>
7 April	<u>DEADLINE: Submit final dissertation</u> 12 noon

„Genozid“. Grenzen und Möglichkeiten eines Forschungsbegriffs - ein Literaturbericht

Der Begriff „Genozid“ ist eine moderne Erfindung. Er ist ein Kunstwort, das sich aus dem griechischen *genos* und einer Ableitung des lateinischen *caedere* (töten) zusammensetzt. Es wurde 1944 durch den Rechtsanwalt Raphael Lemkin (1900-1959) geprägt.¹ In erster Linie vom planmäßigen Massenmord an den europäischen Juden, aber auch schon von der 1915 verübten Ermordung zahlreicher Armenier durch die türkische Regierung abgeleitet, bestritt Lemkin, dass die völkerrechtliche Souveränität eines Staates das Recht einschließen dürfe, Millionen unschuldiger Menschen zu töten.² Nachdem am 8. August 1945 im Londoner Abkommen ein Prozess gegen die deutschen Hauptkriegsverbrecher beschlossen worden war, trug dieser in Polen geborene Jurist, der nach seiner Flucht in die USA im Jahre 1939 an verschiedenen dortigen Universitäten lehrte, in den Folgejahren maßgeblich zur Verankerung des völkerrechtlichen Straftatbestands „Genozid“ in der Völkermordkonvention der Vereinten Nationen bei.

Seitdem hat der Begriff eine beispiellose Karriere gemacht. Im Alltagssprachlichen Diskurs erfährt er gegenwärtig eine geradezu inflationäre Anwendung; internationale Akteure sowie Medien verwenden ihn mit großer Selbstverständlichkeit. So sprechen militante Tierschützer inzwischen von einem „Genozid an Tieren“ bzw. dem „Holocaust auf dem Teller“,⁴ und auch in anderen Zusammenhängen wie Drogenmissbrauch oder Geburtenkontrolle wird der Begriff „Genozid“ bemüht, um politischen Forderungen Gehör zu verschaffen.⁵ „Genozid“ ist somit zu einem hochpolitisierten Kampfbegriff geraten, dessen Drastik politische Forderungen gegen alle Anwürfe immunisieren soll.

Im eigentlichen Kernbereich, dem Völkerrecht, hat er erst seit zehn bis fünfzehn Jahren verstärkte Aufmerksamkeit gefunden, wie William A. Schabas in einer umfangreichen und differenzierten Studie zur völkerrechtlichen Geschichte und Definition des Völkermordes aufzeigen konnte.⁶ Über die Jahrzehnte des „Kalten Krieges“ und seiner dualen Blockstruktur sei die Konvention in der völkerrechtlichen Praxis nahezu bedeutungslos geblieben, wohl auch weil die Akte von „mass violence“ fehlten, die sie im Auge hatte. Jedoch blieben

Raphael Lemkin, *Axis Rule in Occupied Europe. Laws of Occupation, Analysis of Government, Proposals for Redress*. Washington 1944.

- 2 Vgl. Dominik J. Schaller, *Genozidforschung, Begriffe und Debatten*, in: Ders./Rupen Boyadjan u. a. (Hrsg.), *Enteignet. Vertrieben. Ermordet. Beiträge zur Genozidforschung*. Zurich 2004, S. 9-26, hier S. 10.
- 3 Boris Barth: *Genozid. Völkermord im 20. Jahrhundert. Geschichte. Theorien. Kontroversen*, München 2006, S. 8. Vgl. Frank Chalk, *Redefining Genocide*, in: George J. Andreopoulos (Hrsg.), *Genocide. Conceptual and Historical Dimensions*. Philadelphia 1994, S. 47-63, hier S. 47.
- 4 Im März 2003 zog die Tierschutzorganisation PETA mit diesem Slogan die Kritik jüdischer Gruppen auf sich. Vgl. Schaller, *Genozidforschung*, (Anm.2), S. 15, sowie Jürgen Zimmerer: *Kolonialer Genozid? Vom Nutzen und Nachteil einer historischen Kategorie für eine Globalgeschichte des Völkermordes*, in: Schaller/Boyadjan u. a. (Hrsg.): *Enteignet. Vertrieben. Ermordet*, (Anm.2), S. 109-128, hier S. 110.
- 5 Helen Fein, *Genocide, Life Integrity, and War Crimes. The Case for Discrimination*, in: George J. Andreopoulos (Hrsg.), *Genocide*. Philadelphia 1994, S. 95-107, hier S. 95. Vgl. Barth, *Genozid*, (Anm.3), S. 29.
- 6 William A. Schabas: *Genozid im Völkerrecht*. Aus dem Engl. v. Holger Fliessbach. 792 S., Hamburger Edition, Hamburg 2003 (Orig.-Ausg. Cambridge 2000).

die Verbrechen in Ost-Timor, Kambodscha, Burundi und Äthiopien in den 70er Jahren unbeachtet.⁷ Erst seit dem Zusammenbruch der meisten sozialistischen Staaten in den 1990er Jahren und aus dem Prokrustesbett des Systemgegensatzes entlassen habe sich seit Beginn der 90er Jahre eine neue Sensibilität gegenüber Menschenrechtsverletzungen bewiesen.⁸ Die Einrichtung der Ad-hoc-Strafgerichtshöfe für das ehemalige Jugoslawien (ICTY) und Ruanda (ICTR) sowie schließlich 2002 des Internationalen Strafgerichtshofes (ICC) zeugen davon.⁹ Gleichzeitig reflektieren diese neuen Rechtsinstitute auf Gewalttaten, die im Zuge der „neuen Kriege“¹⁰ und ihrer mörderischen Logik des *bellum se ipse alet* mit der Genozidkonvention zu erfassen waren.

Nach wie vor ist die Hauptfrage, wie das nationalstaatliche Souveränitätsprinzip in Fällen genozidaler Gewalt eingeschränkt werden kann. Die Konvention gibt darauf eine unbefriedigende Antwort, überlässt sie doch in Artikel IV die Bestrafung von Völkermorden absurderweise dem genozidalen Staat.¹¹ Daneben bleibt unklar, wie Artikel I der Konvention zu verstehen ist. Hiemach verpflichten sich die unterzeichnenden Staaten, Völkermord zu verhüten. Berechtigt diese Klausel die internationale Gemeinschaft zur „humanitären Intervention“, verpflichtet sie diese sogar?¹² Schabas kann nachweisen, dass humanitäre Verpflichtungen zwar gem angeführt werden, sofern dabei politische Interessen verfolgt werden können wie im Kosovo. Sind diese Interessen aber nicht vorhanden oder wenig ausgeprägt, wie etwa in diversen Krisenregionen Afrikas, bleibt die Bereitschaft zum Engagement trotz aller Beteuerungen, die Versäumnisse von Ruanda nicht zu wiederholen, gering.¹³ *Samantha Power* schlägt als Konsequenz dieser Unzulänglichkeiten eine unilaterale Doktrin der humanitären Intervention vor.¹⁴ Dieser Automatismus wird jedoch als völkerrechtlich problematisch und politisch keineswegs immer klug angesehen. Außerdem scheinen Militärinterventionen zwar alte Übel zu beseitigen, aber auch neue Gewalttaten zu generieren, insbesondere wenn es an der Planung für die Zeit *nach* dem Militäreinsatz mangelt.¹⁵ Diese Diskussionen sind nicht abgeschlossen und werden uns auch im 21. Jahrhundert weiter verfolgen. Ermutigende Erfolge lassen sich in der völkerrechtlichen Ahndung von Verbrechen gegen die Menschheit, Kriegsverbrechen und Genozid aufzählen. Das Jugoslawien-

7 Ein möglicher Anwendungsfall wäre der Prozess gegen Adolf Eichmann (April bis Dezember 1961) gewesen, den der Staat Israel jedoch vor einem nationalen, nicht vor einem internationalen Gericht führte. Vgl. Schabas, Genozid im Völkerrecht, (Anm.6), S. 136.

8 Vgl. auch Kofi Annan, Verhütung von Völkermord, in: Deutsches Institut für Menschenrechte (Hrsg.), Jahrbuch für Menschenrechte 2005. Frankfurt am Main 2004, S. 259-262.

9 Vgl. nur <<http://www.icc-cpi.int>> [Zugriff am 10.02.2007]. Bei der Einrichtung des Internationalen Gerichtshofs in Den Haag griff man zur Definition des Verbrechens „Genozid“ (Art. 6 ICC-Statut) mangels Alternativen wortwörtlich auf Art. II der Konvention von 1948 zurück. Barbara Luders, Die Strafbarkeit von Völkermord nach dem Römischen Statut für den Internationalen Strafgerichtshof. Berlin 2004, S. 14.

10 Kaldor, Mary, Neue und alte Kriege, Frankfurt/M. 2000; Herfried Munkler, Die neuen Kriege, Hamburg 2002.

11 EIC; in Kürsat-Ahlens: Ober das Täten in Genoziden. Eine Bilanz historisch-soziologischer Deutungen, in: Peter Gleichmann/Thomas Kuhne (Hrsg.), Massenhaftes Täten. Kriege und Genozide im 20. Jahrhundert. Essen 2004, S. 180-206, hier S. 203.

12 Vgl. auch Michael Walzer, Mehr als humanitäre Intervention. Menschenrechte in der globalen Gesellschaft, in: Internationale Politik 60 (2005), Nr. 2, S. 8-20.

13 Vgl. Samantha Power: A Problem from Hell. America and the Age of Genocide. 512 S., Basic Books, New York 2002, S. 329ff.

14 Ebd., S. 502ff.

15 Anson Rabinbach, Lemkins Schöpfung. Wie Völkermord zum juristischen und politischen Begriff wurde, in: Internationale Politik 60 (2005), Heft 2, S. 21-31, hier S. 21.

Tribunal schuf Präzedenzfälle, indem es Vergewaltigungen in die Liste der Kriegsverbrechen aufnahm. Massive Menschenrechtsverletzungen im Sudan und nun auch in Kambodscha – 30 Jahre nach der Schreckensherrschaft der Roten Khmer – werden strafrechtlich verfolgt. Die neue liberianische Präsidentin Ellen Johnson-Sirleaf hat die Auslieferung ihres Amtsvorgängers Charles Taylor beantragt, der vor dem Special Court for Sierra Leone (SCSL) zusammen mit 12 weiteren Hauptverantwortlichen der Kriegsverbrechen angeklagt ist und im nigerianischen Exil sitzt. Vor diesem Gericht, das auf einer neuen Variante internationaler Strafgerichtsbarkeit beruht, nämlich auf einer Vereinbarung zwischen der UN und der Regierung von Sierra Leone, wird zum ersten Mal die Rekrutierung von Kindersoldaten als Kriegsverbrechen verhandelt.¹⁶ Ob allerdings einer der brutalsten Warlords Afrikas, Joseph Kony, Führer der *Lord's Resistance Army* in Uganda vor den ICC kommt, bleibt trotz internationalen Haftbefehls ungewiss.¹⁷

Fällt das Resümee unter Juristen in Bezug auf die Prävention von Genoziden pessimistisch und verhalten positiv im Hinblick auf die Aburteilung von Tätern *post festum* aus, so ist der Begriff und mit ihm seine Typologie in der Wissenschaft deutlich umstrittener. Festzuhalten ist zunächst, dass der Terminus in den letzten Jahren als Kategorie der Analyse von Massenverbrechen starke Resonanz gefunden hat. Mittlerweile existieren diverse Zeitschriften¹⁸, Institute¹⁹, Enzyklopädien²⁰, es gibt ein European Network of Genocide Research, und immer größer wird der Berg an Fallstudien sowie Gesamtdarstellungen²¹. Im Laufe des Jahres 2007 soll eine online *Encyclopedia of Mass Violence* ins Netz gestellt werden.²² Diese Entwicklung ist ein Reflex auf die zunehmend hoffnungslos unübersichtli-

16 Vgl. z. B. die Konferenz: Crimes against Humanitarian Law, International Trials in Perspective, 24.02.2006, <<http://www.tcd.ie>> [Zugriff am 10.02.2007]; Andrea Bähm, Der Nächste, bitte! Die Zeit vom 05.04.2006.

17 Mareike Schomerus, Die Geißel seines Volkes, in: Die Zeit vom 13.07.2006.

18 Z. B. Holocaust and Genocide Studies; Journal of Genocide Research; Zeitschrift für Genozidforschung.

19 Z. B. Institut für Diaspora- und Genozidforschung, Universität Bochum; Center for the Prevention of Genocide (CPG) (Netzwerk); Genocide Watch der Leo Kuper Foundation; Genocide Studies Program am Yale Center for International and Area Studies; Center for Holocaust and Genocide Studies, Clark University.

20 Z. B. Encyclopedia of genocide and crimes against humanity, Dinah L. Shelton, ed. in chief, Detroit u. a. 2005; Encyclopedia of genocide, Israel W. Charny, ed. in chief, Santa Barbara 1999; Gunnar Heinsohn (Hrsg.), Lexikon der Völkermorde, Reinbek 1998.

21 Irmlud Wojak/Susanne Mehl (Hrsg.): Völkermord und Kriegsverbrechen in der ersten Hälfte des 20. Jahrhunderts. Frankfurt am Main, New York 2004; Bruce Wilshire: Get 'em all, kill 'em! Genocide, Terrorism, Righteous Communities. New York, Toronto u. a. 2005; Eric D. Weitz: A Century of Genocide. Utopias of Race and Nation. Princeton 2003; Benjamin A. Valentino: Final Solutions. Mass killing and Genocide in the twentieth Century. Cornell University Press, Ithaca 2004; Samuel Totten: (Hrsg.): Genocide at the Millennium. New Brunswick, London 2005; Colin Tatz: With Intent to Destroy. Reflecting on Genocide. London, New York 2003; Martin Shaw: War and Genocide. Organized Killing in Modern Society. Cambridge 2003; Jacques Semelin, Purifier et Détruire. Usages politiques des massacres et génocides, Paris 2005; William D. Rubenstein: Genocide. A History. Harlow 2004; Verena Radkau u. a. (Hrsg.): Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert. Wien u. a. 2004; Manus I. Midlarsky: The Killing Trap. Genocide in the Twentieth Century. Cambridge 2005; Adam Jones (Hrsg.): Völkermord, Kriegsverbrechen und der Westen. 534 S., Parthas, Berlin 2005; Robert Gellately/Ben Kiernan (Hrsg.), The Specter of Genocide. Mass Murder in Historical Perspective, Cambridge 2003; Robert S. Frey (Hrsg.), The Genocidal Temptation, Auschwitz, Hiroshima, Rwanda and Beyond, Lanham 2004; Omer Bartov: Mirrors of Destruction. War, Genocide, and Modern Identity. Oxford 2000; Barth, Genozid (Anm.3); Alex Alvarez: Governments, Citizens, and Genocide. A Comparative and Interdisciplinary Approach. Bloomington 2001.

22 <<http://www.ceri-sciencespo.com/geno/index.htm>> [Zugriff am 10.02.2007].

che Gewaltwelt gerade des 20., mutmaßlich auch des 21. Jahrhunderts. Viele Disziplinen versuchen sich an Systematisierungen, die internationalen Rechtswissenschaften, die Historiographie, Politologie, Soziologie, Psychologie und Anthropologie tragen jeweils unterschiedliche Forschungsschwerpunkte an den Terminus heran. Und in der Tat wissen wir heute sehr viel mehr über Handlungslogiken, Gelegenheiten und Kontexte kollektiver Gewalt als noch vor zwanzig Jahren. Von einer allseits akzeptierten Theorie des „Genozids“, die sowohl zur Beschreibung als auch zur Erklärung Substantielles beitragen und zudem auch juristisch befriedigen konnte, kann indessen keine Rede sein.²³ Die Meinungen, was ein Genozid sei, welche strukturellen Regelmäßigkeiten diesen charakterisieren, ob man das Phänomen allein auf die Moderne begrenzen sollte, wie man es von ethnischen Säuberungen, Massakern oder Terror abgrenzen oder ob die Shoah das entscheidende beispielgebende Muster für diese Analyse sein sollte, gehen weit auseinander. Dementsprechend lautet das Fazit des Historikers an der Princeton University, Anson Rabinbach:

„Die Hauptschwierigkeit besteht für alle Theoretiker darin, dass der Versuch, einen einzigen ‚generischen Begriff‘ zu finden, der alle Formen vergangener, gegenwärtiger und künftiger Genozide umfasst, nur Definitionen hervorbringt, denen es an Klarheit und Substanz mangelt und die entweder zu elastisch oder zu spezifisch sind.“²⁴

Schon die Genese des Begriffs verweist auf seine zentrale Problematik, nämlich die enge Verknüpfung von ethischen, politischen, wissenschaftlichen und juristischen Dimensionen; eine Verschmelzung, die bis heute andauert und die dazu führt, dass der Begriff viele, vielleicht zu viele Anforderungen erfüllen soll:

- zum ersten soll er sehr heterogene historische Phänomene auf einen Nenner bringen
- an diese Gemeinsamkeiten anknüpfend soll er zum zweiten systematische Vergleiche von Massenverbrechen ermöglichen
- schließlich soll er drittens auch noch zu einer Art internationalem Frühwarnsystem für zukünftige Genozide ausgebaut werden – und damit die regelmäßig zu beobachtende Tatenlosigkeit der internationalen Gemeinschaft aufbrechen²⁵

23 Vgl. Barth, Genozid, (Anm.3), S. 7. Verschiedene Autoren neigen dazu, den Begriff vorbehaltlich einer endgültigen Klärung als „schweres Verbrechen gegen die Menschlichkeit“ zu bezeichnen; vgl. Adam Jones, Einführung. Mittäterschaft und Geschichte, in: Ders. (Hrsg.), Völkermord, Kriegsverbrechen und der Westen. Berlin 2005, S. 11-40, hier S. 36. Auch Yves Temon warnt vor einem inflationären Gebrauch des Wortes „Völkermord“, da es emotional zu stark besetzt sei; vgl. Ders., Perzeption und Prävention des Genozid, in: Verena Radkau/Eduard Fuchs u. a. (Hrsg.), Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert. Wien 2004, S. 16-21, hier S. 17. Gerlach lehnt den Begriff als sozialwissenschaftliche Kategorie ab, vgl. Christian Gerlach, Nationsbildung im Krieg, Wirtschaftliche Faktoren bei der Vernichtung der Armenier und beim Mord an den ungarischen Juden, in: Hans-Lukas Kieser/Dominik J Schaller (Hrsg.): Der Völkermord an den Armeniern und die Shoah. 648 S., Chronos Verlag, Zürich 2002, S. 347-422, hier S. 350. Dirk Moses wiederum plädiert mittlerweile für „Critical Genocide Studies“; vgl. seinen Beitrag „Historiography of Comparison: Comparative or Critical Genocide Studies?“ auf der Tagung Genocides: Forms, Causes and Consequences. The Namibian War (1904–08) in historical perspective, vgl. den Tagungsbericht auf <<http://www.hist.net/ag-genozid/tagungsbericht.htm>> [Zugriff am 10.02.2007].

24 Anson Rabinbach, Lemkins Schöpfung. Wie Völkermord zum juristischen und politischen Begriff wurde, in: Internationale Politik 60 (2005), Heft 2, S. 21-31, hier S. 24; vgl. auch Barth, Genozid, (Anm.3), S. 7f.

25 Ted R. Gurr (Hrsg.), Minorities at Risk, A Global View of Ethnopolitical Conflicts. Washington 1993; ders./John L. Davies (Hrsg.), Preventive Measures, Building Risk Assessment and Crisis Early Warning Systems. Lanham 1998; Annan, Verhütung, (Anm.8), S. 259-262. Vgl. auch Gregory Stanton, Wie wir Genozid verhindern können. Der Aufbau einer internationalen Kampagne zur Beendigung von Genoziden, in: Radkau u. a. (Hrsg.), Genozide, (Anm.21), S. 29-39; auch Power, Problem from Hell, (Anm.13); Samuel Totten, The Role of Nongovernmental Organizations in Addressing the Prevention,

- viertens interessieren sich die Staatsrechtler v. a. fOr die Tatdefinition und deren Sanktionierung

- zu diesen voneinander abweichenden Erkenntnisinteressen tritt fOnftens noch die moralische Aufladung des Begriffs, gerade in der allenthalben zu lesenden Formel „crime of crimes“. Diese Wertung legt eine Skala nahe, der die fatale „Konkurrenz der Opfer“ auf dem FuBe folgen muss.²⁶ Zwar gibt es aus juristischer Warte keine Hierarchie zwischen verschiedenen Verbrechen gegen die Menschheit.²⁷ Dennoch schwingt diese insbesondere im medialen Interesse immer mit. Und so ist das Anliegen von Hinterbliebenen und Opfern, die ihre Schicksale nicht an als Genozid klassifizierte organisierte exzessive Gewaltakte binden können, „angemessen“ berücksichtigt zu werden und ihrem Leiden Gehör zu verschaffen, verständlich und nachvollziehbar. Die Genozidforschung steht damit vor dem unauflösbaren Dilemma zwischen einer distanzierten Wissenschaftlichkeit, die sich der Problematik einer klassifizierenden Betrachtung stellen muss, und einer Rücksicht nehmenden Sensibilität, die Gefühle nicht durch eine Relativierung von Gewalttaten verletzen möchte.²⁸

Folge wie Ursache der Definitionswehen und der anspruchsgeladenen Aufladung des Begriffs ist, dass Teile der Forschung dazu neigen, in einem Genozidkonzept sämtliche historischen Fälle von *mass violence*, *mass killings*, *mass murder*, *mass atrocities*, *mass crimes* abdecken zu wollen.²⁹ Aber weisen der Mord an 7.000 bosnischen Männern 1995 in Srebrenica durch serbisch-nationalistische Milizen, der Völkermord an Tutsi, Tutsi-Mischlingen und gemäßigten Hutu, dem 1994 in Ruanda innerhalb von 13 Wochen über 800.000

Intervention, and Punishment of Genocide in the 1980s, 1990s, and Early 2000s, in: Ders. (Hrsg.), *Genocide at the Millennium*, (Anm.21), S. 75-111. Skeptisch dagegen: Thomas Cushman, Is genocide preventable? Some theoretical considerations, in: *Journal of Genocide Research* 5 (2003), Heft 4, S. 523-542, sowie Alan Kramer, *The War of Atrocities. Murderous Strategies and Extreme Combat*, in: Alf LÖdtker (Hrsg.), *No Man's Land of Violence. Extreme Wars in the 20th Century*. Göttingen 2006, S. 11-33. Jones zufolge wäre eine zukunftsweisende Typologie aller vorstellbaren Genozide eine „Meisterleistung der Fantasie“. Jones, *Völkermord*, (Anm.23), S. 42.

26 Vgl. Schabas, *Genozid im Völkerrecht*, (Anm.6), S. 9, S. 91f.; Jean-Michel Chaumont: *Die Konkurrenz der Opfer. Genozid, Identität und Anerkennung*. 359 S., Zu Klampen, Lüneburg 2001; auch Yves Ternon, *Perzeption und Prävention des Genozids*, in: Radkau u. a. (Hrsg.), *Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert*, (Anm.21), S. 16-21, hier S. 17f.

27 Schabas, *Genozid im Völkerrecht*, (Anm.6), S. 117f.

28 Barth, *Genozid*, (Anm.3), S. 53; vgl. auch Tatz, *With Intent to Destroy*, (Anm.21), S. 1-16, 19f.; Radkau, Verena, *Einleitung*, in: dies. u. a. (Hrsg.), *Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert*, (Anm.21), S. 9-14, hier S. 11; Miran Dabag, *Genozidforschung. Leitfragen, Kontroversen, Oberlieferung*, in: *Zeitschrift für Genozidforschung* 1 (1999), S. 6-35, hier S. 23; Eric D. Weitz, *Holocaust, Genozid und die Macht der Definition*, in: Verena Radkau u. a. (Hrsg.), *Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert*, (Anm.21), S. 52-59, hier S. 54.

29 Vgl. Frank Chalk/Kurt Jonassohn: *Genozid – ein historischer Überblick*, in: Miran Dabag/Kristin Platt (Hrsg.): *Genozid und Moderne. Bd. 1, Strukturen kollektiver Gewalt im 20. Jahrhundert*. Leske und Budrich, Opladen 1998, S. 294-308; auch Alexander Hinton, *The Dark Side of Modernity*, in: Ders. (Hrsg.), *Annihilating Difference. The Anthropology of Genocide*. Berkeley, Los Angeles 2002, S. 1-40, hier S. 7. Vgl. auch die Beispiele in Mark Levene, *Warum ist das 20. Jahrhundert das Jahrhundert der Genozide?* In: *Zeitschrift für Weltgeschichte* 5 (2004), Heft 2, S. 9-37, hier S. 9, sowie Samuel Totten/William S. Parsons/Israel Charny, *Confronting Genocide and Ethnocide of Indigenous Peoples. An Interdisciplinary Approach to Definition, Intervention, Prevention, and Advocacy*, in: Alexander L. Hinton (Hrsg.), *Annihilating Difference*. Berkeley 2002, S. 54-91, hier S. 67. Anschauliches Beispiel für die Spannung, unter der die Genozidforschung leidet, ist die irritierende Sicherheit, mit der Wolfgang Benz, *Vermeidung der Wahrheit. Der Bundestag, die Türkei, der Genozid an den Armeniern – und von der Verfälschung der Geschichte durch falsche Worte*, in: *Frankfurter Rundschau* vom 12.07.2005, S. 7, den Terminus verteidigt. Um *politically* und *morally correct* zu sein, werden alle Ambivalenzen und Inkonsistenzen des Ansatzes weggedeckt.

Menschen zum Opfer fielen, der sich Liber drei Jahre hinziehende Holocaust und vier koloniale counter-guerilla-Feldzug gegen die Herero 1904-1908, dem mehrere zehntausend Afrikaner zum Opfer fielen, solche Gemeinsamkeiten auf, dass es sinnvoll ist, alle diese Ereignisse als „Genozid“ zu deklarieren? Oder sind es nicht doch sehr unterschiedliche Phänomene organisierter Gewaltentgrenzung? Andersherum gefragt: Versperren nicht das Festhalten an diesem Terminus und den Vorgaben und Vorannahmen, die die Konvention macht und an denen sich alle Wissenschaftlerinnen und Wissenschaftler mitunter nolens volens abarbeiten, Einsichten Liber das empirisch höchst vielgestaltige Wesen von entgrenzter Gewalt? Taugt vier Begriff noch als heuristisches Hilfsmittel, die Vielfalt von exzessiver Gewalt im 20. Jahrhundert zu verstehen? Es scheint an der Zeit, eine Bilanz zu ziehen, was vier Genozid-Terminus an Erklärungen leisten kann und wo seine Grenzen liegen. Im Folgenden sollen dazu in einem ersten Teil die sich aus der UN-Konvention ergebenden Probleme für eine vergleichende Genozidforschung diskutiert werden. In einem zweiten Abschnitt werden am Beispiel einzelner Aspekte Ergebnisse und offene Fragen vorgestellt. Grundlage des Beitrags bilden neuere Veröffentlichungen der vergleichenden Genozidforschung sowie einige aktuelle Diskussionsbeiträge aus deutschen Printmedien.

I. Die Probleme

1. Die UN-Konvention

Die juristische Definition in der von den Vereinten Nationen am 9. Dezember 1948 verabschiedeten „Konvention Liber die Verhütung und Bestrafung des Völkermords“ legte den Begriff sehr weit aus und umfasste entgegen vier deutschen Bezeichnung keineswegs nur Mord:

„In dieser Konvention bedeutet Völkermord eine vier folgenden Handlungen, die in der Absicht begangen werden, eine nationale, ethnische, rassische oder religiöse Gruppe als solche ganz oder teilweise zu zerstören:

- a) Täten von Mitgliedern vier Gruppe;
- b) Verursachung von schwerem körperlichem oder seelischem Schaden an Mitgliedern vier Gruppe;
- c) Vorsätzliche Auferlegung von Lebensbedingungen für die Gruppe, die geeignet sind, ihre körperliche Zerstörung ganz oder teilweise herbeizuführen;
- d) Verhängung von Maßnahmen, die auf die Geburtenverhinderung innerhalb der Gruppe gerichtet sind;
- e) Gewaltsame Überführung von Kindern der Gruppe in eine andere Gruppe.“³⁰

Drei Fixpunkte bestimmen die Konvention: die Festlegung auf eine *Gruppe* als Opfer, auf die *Intention* als handlungsleitend für die Täter und auf die (*partielle*) *Zerstörung* als Tat. Schon früh wurde vier Vorwurf geäußert, die Definition der zu schützenden Gruppen sei zu eng, die der Taten zu weit.³¹ Bei der Formulierung der Konvention richtete man sich offen-

30 Zit. nach Schabas, Genozid im Völkerrecht, (Anm.6), S. 721. Zu einer genauen Analyse der Prämissen in der Definition vgl. auch ebd., S. 139-201.

31 Eric D. Weitz, Holocaust, Genozid und die Macht der Definition, in: Radkau u. a. (Hrsg.), (Anm.21), S. 54. Vgl. Ternon, Perzeption und Prävention, (Anm.26), Ebd., S. 16, sowie Tatz, Intent to Destroy, (Anm.21), S. 73, und Schaller, Genozidforschung, (Anm.2), S. 12.

bar nach einem objektiven Gruppenbegriff.³² Nicht die Selbstwahrnehmung, sondern eine spezifische nationale etc. Struktur bestimmt den Gruppenbegriff der Konvention. Auf Wunsch der Sowjetunion wurden politische Gruppen, anders als von Lemkin ursprünglich intendiert, nicht in die Konvention aufgenommen. Der kulturelle Genozid, mit dem Raphael Lemkin die massive Unterdrückung einer Sprache sowie die Zerstörung religiöser wie kultureller Denkmäler und Institutionen erfasst wissen wollte, wurde vor allem auf Drängen der USA, Frankreichs und der Niederlande als ein mögliches Völkermordverbrechen fallen gelassen. Sie befürchteten als ehemalige Kolonialmächte nachträglich in Genozidverdacht zu geraten.³³ Dieser Eingrenzung in der Benennung der Opfergruppen steht eine sehr weite Auslegung der als genozidal qualifizierten verbrecherischen Praktiken gegenüber. Augenscheinlich war der Wille maßgeblich, angesichts der Massenverbrechen in der ersten Hälfte des 20. Jahrhunderts ein möglichst breites Spektrum an Tatbeständen in die Konvention aufzunehmen.

Das hat die verstörende Konsequenz, dass nicht jede massive Menschenrechtsverletzung als „Genozid“ qualifiziert wird, während mitunter weniger grausam scheinende Vergehen genozidale Akte darstellen sollen. Brutale Verbrechen, wie sie etwa in Südafrika der Apartheid praktiziert wurden, werden nicht als Genozid klassifiziert, weil – so Colin Tatz –, die Vernichtungsabsicht nicht nachzuweisen sei. Diese Art der Gewalt sollte nicht zerstören, sondern disziplinieren. Das Apartheid-Regime hatte die Schwarzen in einer ständigen Untertanenstellung halten wollen. Dagegen stelle die Überführung australischer Aborigine-Kinder in staatliche Schulheime einen Völkermord dar, weil sie die ethnische Identität der Kinder tilgte.³⁴ Oder, um ein anderes Beispiel zu nennen: Die Völkermordkonvention verurteilt sinngemäß die Aushungerung von zwei ukrainischen Bauern als Ukrainer, nicht aber die Aushungerung von zwei Millionen Bauern in der Ukraine, denn Bauern- oder Klassenmord ist kein Völkermord.³⁵ Auch bei den Gewalttaten an den Armeniern 1915 stellt *Christian Gerlach* die Frage, ob hier tatsächlich eine ethnische Gruppe zerstört werden sollte oder ob nicht ökonomische Erwägungen einen wichtigen Grund für die Massaker darstellen.³⁶ Diese Morde waren damit nicht legal, sie stellten Verbrechen gegen die Menschheit dar, waren aber trotz der immens hohen Zahlen und der vergleichsweise großen Homogenität der Opfergruppe kein Völkermord. Ähnlich ist es auch im Fall der Ereignisse in Kambodscha. Hier scheint die Konstatierung von „Gruppen“ wenig sinnvoll, da Täter und Opfer überwiegend Angehörige einer Gruppe waren bzw. aus mehreren Gruppen stammten. Selbst wenn innerhalb der verschiedenen ineinander greifenden Mordaktionen der Roten Khmer auch eth-

32 Vgl. Barbara Luders, Strafbarkeit von Genoziden nach dem Römischen Statut, (Anm.9), S. 51; Schaller, Genozidforschung, (Anm.2), S. 14; Schabas, Genozid im Völkerrecht, (Anm.6), S. 148ff.; Barth, Genozid, (Anm.3), S. 23; Elçin Kurşat-Ahlers, Über das Töten in Genoziden, (Anm.11), S. 191; Frank Chalk/Kurt Jonassohn, Genozid – ein historischer Überblick, (Anm.29), S. 300.

33 Gerd Hankel, Was heißt eigentlich Völkermord? Überlegungen zu einem problematischen Begriff, in: *Mittelweg* 36, 14. Jg., 2005, Heft 4, S. 70-81; Schabas, Genozid im Völkerrecht, (Anm.6), S. 237-251.

34 Colin Tatz, With intent to destroy, (Anm.21), zu Südafrika S. 107ff., zu Australien S. 67ff. Vgl. auch Barth, Genozid, (Anm.3), S. 31-32; Dirk A. Moses (Hrsg.), *Genocide and Settler Society. Frontier Violence and Stolen Indigenous Children in Australian History*, Oxford 2004.

35 Zu der umstrittenen Einordnung der Millionen von Hunberttoten in der Ukraine, dem Holodomor, vgl. Barth, Genozid, (Anm.3), S. 136-145.

36 Christian Gerlach: Wirtschaftliche Faktoren bei der Vernichtung der Armenier und beim Mord an den ungarischen Juden. in: Hans-Lukas Kieser/Dominik J. Schaller (Hrsg.): *Der Völkermord an den Armeniern und die Shoah*. 648 S., Chronos Verlag, Zürich 2002, S. 347-422.

nische und religiöse Gruppen getötet wurden,³⁷ waren die meisten Opfer des Pol-Pot-Regimes Khmer.³⁸ Die Erfindung des Gegners erfolgte also über andere Operationen als die einer Gruppenzuordnung. Die von Jean Lacouture geprägte Bezeichnung „Auto-Genozid“ ist aber problematisch, weil sie die Grenze zwischen Tätern und Opfern verwischt.³⁹ Zudem ist nach wie vor ungeklärt, ob und wann die Gruppenzugehörigkeit der Opfer überhaupt für ihre Ermordung relevant wurde. Mit diesem womöglich fehlenden Kriterium der Zuordnung zu einer Gruppe durch die Täter wird aber die Klassifikation als Genozid nach der Völkermordkonvention fraglich.⁴⁰ Nicht nur, dass im Fall dieses asiatischen Staates die Gruppenlogik der Konvention nicht richtig greift, auch die Definition der Opfer verkompliziert den Fall. Wenn die Morde von Pol Pots Steinzeitsozialisten einen politischen Charakter getragen haben, können sie nach der UN-Definition kein Genozid gewesen sein. Manche Forscher bezeichnen diese Schlussfolgerung angesichts der hohen Todesraten wiederum als „paradox“ und weisen sie zurück.⁴¹

Die vorgeschlagenen Lösungen aus diesem Ungleichgewicht der juristischen Bestimmungen weisen fast alle in die gleiche Richtung: Es wird der Wunsch geäußert, wissenschaftlich kreativ vorzugehen und neue Sub-Typologien zu bilden, die von Begriffen wie „Ethnozid“, „Demozid“, „Ökonomizid“, „Klassizid“, „Politizid“, „Femizid“, „Gendercide“ bis zu Kompromiss-Konzepten wie „genozidalen Massakern“ reichen.⁴² So spricht sich z. B. *Jean-Michel Chaumont* dafür aus, das Phänomen „Ethnozid“ als *Verdrängung* einer Kultur differenzierter vom Phänomen „Genozid“ als kollektivem *Massenmord* zu unterscheiden.⁴³ R. J. Rummel mochte unter „Politizid“ sowohl Hitlers „Säuberung“ der SA im Jahre 1934 wie lybische Bombenattentate auf zivile Passagiermaschinen fassen.⁴⁴ Wie man allerdings alle

37 So etwa die Vietnamesen, die buddhistischen Mönche und die muslimischen Cham; vgl. John D. Ciorciari, „Auto-Genocide“ and the Cambodian Reign of Terror, in: Schaller/Boydzian u. a. (Hrsg.), *Enteignet. Vertrieben. Ermordet*, (Anm.2), S. 413-435, hier S. 423ff.

38 Ciorciari, „Auto-Genocide“, (Anm.37), S. 414.

39 Barth, *Genozid*, (Anm.3), S. 149.

40 Hans-Joachim Heintze, *Die Verbrechen der Roten Khmer und die Völkermord-Konvention*, in: *Zeitschrift für Genozidforschung* (2004), Nr. 2, S. 114-126; Micha Brumlik, *Zu einer Theorie des Völkermords*, in: *Blätter für deutsche und internationale Politik* 8 (2004), S. 923-932, hier S. 926.

41 Vorsichtige Schätzungen gehen davon aus, dass in Kambodscha während der Herrschaft der Roten Khmer ca. 21 Prozent der Gesamtbevölkerung getötet wurden; vgl. Hurst Hannum/David Hawk, *The case against the standing committee of the communist party of Kampuchea*. New York 1986, S. 82-138. Vgl. Ciorciari, „Auto-Genocide“, (Anm.37), S. 428; vgl. auch Leo Kuper, *The Prevention of Genocide*. New Haven 1985; Ben Kieman, *The Pol Pot Regime. Race, Power and Genocide in Cambodia under the Khmer Rouge, 1975-1979*, Yale 2002.

42 Vgl. Peter Imbusch, *Probleme der deutschen Genozidforschung. Eine Übersicht*, in: *Mittelweg* 36, 10 (2001), Heft 2, S. 49-53; vgl. auch die Einträge in Heinsohn (Hrsg.), *Lexikon* (Anm.20).

43 Jean-Michel Chaumont: *Die Konkurrenz der Opfer. Genozid, Identität und Anerkennung*. 359 S., Zu Klampen, Luneburg 2001, S. 183ff.

44 I.W. Charny (Hrsg.), *Encyclopedia of Genocide*, Bd. 1, S. 22. Vahakn N. Dadrian, *A Typology of Genocide*, in: *International Review of Modern Sociology* 5 (1975), S. 201-212, unterscheidet nach der Motivation des Täters fünf Genozid-Typen:

1. Kulturelle Genozide, deren Ziel es ist, die Opfergruppe in die eigene Gruppe zu assimilieren;
2. Latente Genozide, die sich häufig ohne gezielte Maßnahmen des Täters, wie etwa durch die Verbreitung von Krankheiten oder die Inkaufnahme von Hunger vollziehen;
3. Retributive Genozide, deren Motivation in Strafmaßnahmen gegen Aufständische oder Minderheiten besteht;
4. Utilitaristische Genozide, die in erster Linie verübt werden, um sich die ökonomischen Ressourcen der Opfergruppe anzueignen;
5. Optimale Genozide, deren Bestreben die vollständige Vernichtung einer Zielgruppe ist.

Fälle aufnimmt, ohne dabei den heuristischen Wert des Begriffs Genozid deutlich zu verringern⁴⁵, bleibt ein Geheimnis der „-zider“. *Valentino* schlägt aus diesem Grund eine neue Klassifizierung für alle staatliche massenhafte Gewalt vor: Als Oberbegriff empfiehlt er, den Terminus „Massenmord“ zu wählen. Genozide, politische Massenmorde der kommunistischen Regime und counter-insurgency-Kriege wurden drei eigenständige Unterkategorien bilden.⁴⁶ Als vierten Typus fügt *Barth* die „ethnische Säuberung“ hinzu, die sich in Zielsetzung und Intensität vom Genozid dadurch unterscheidet, dass die Täter die Vertreibung der betroffenen Menschen aus den eigenen Staatsgrenzen anstreben, nicht aber deren Tod.⁴⁷ Daneben existiert eine Minderheit von Autoren, die einem exklusiven Genozidverständnis folgt und den Begriff einengen will, indem sie ihn – den Holocaust vor Augen – tendenziell auf die Absicht bezieht, eine Gruppe *vollständig umzubringen*.⁴⁸ Alle diese Versuche reagieren einerseits auf die empirische Unübersichtlichkeit von kollektivem, entgrenztem Gewalteinsatz. Seit dem 20. Jahrhundert scheinen sich die Grenzen zwischen Pogromen, Vertreibungen, Massakern und Völkermorden immer mehr zu verwischen. Andererseits ist den Neuansätzen mitunter nur implizit, aber dennoch erkennbar das Anliegen zu eigen, nicht nur zu systematisieren, sondern auch zu hierarchisieren, eine Tat für herausgehobener, schwerer, bedeutsamer als die andere zu halten – ein Ansinnen, das menschlich verständlich ist, aber wissenschaftlich begründet werden sollte.

Den nahe liegenden Ansatz, sich von der vorgegebenen Begriffsbestimmung zu lösen und eine Neudefinition für die Forschung zu wagen, wählte die Soziologin Helen Fein:

"Genocide is a series of purposeful actions by a perpetrator(s) to destroy a collectivity through mass or selective murders of group members and suppressing the biological and social reproduction of the collectivity. This can be accomplished through the imposed proscription or restriction of reproduction of group members, increasing infant mortality, and breaking the linkage between socialization of children in the family or group of origin. The perpetrator may represent the state of a victim, another state or another collectivity."⁴⁹

Mit dieser Begriffsbildung bezieht Fein auch die in der UN-Definition ausgesparten politischen und kulturellen Opfergruppen ein. Letztlich aber werde, so ein Einwand, fast jede Opfergruppe aus politischen Gründen bedroht. Zudem seien auch Organisationen wie die SS, der NKWD oder der „Leuchtende Pfad“ in Peru politische Gruppen.⁵⁰ Dagegen werden

Für eine ähnliche, vierstufige Typologie vgl. Chalk/Jonassohn, *The conceptual framework*, (Anm.29), insbes. S. 12-15; anhand der Opfergruppen: Robert Melson, *Problems in the Comparison of the Armenian Genocide and the Holocaust. Definitions, Typologies, Theories, and Fallacies*, in: Stöcker/Färster/Gerhard Hirschfeld (Hrsg.), *Der Genozid in der modernen Geschichte*. Münster 1999, S. 22-35.

45 Vgl. Ben Whitaker, *Revised and Updated Report on the Question of the Prevention and Punishment of the Crime of Genocide*, in: United Nations Economic and Social Council. Commission on Human Rights, E.C.N. 4 Sub. 2, 1985, 6, 2. Juli 1985. Vgl. Hinton, *Dark Side*, (Anm.29), S. 4f.

46 Benjamin A. Valentino: *Final Solutions. Mass killing and Genocide in the twentieth Century*. 317 S., Cornell University Press, Ithaca 2004, S. 9f.

47 Boris Barth: *Genozid. Völkermord im 20. Jahrhundert. Geschichte. Theorien. Kontroversen*. 271 S., Beck, München 2006, S. 190ff.

48 So etwa Steven Katz, *The Holocaust in Historical Perspective*, Bd. 1, *The Holocaust and Mass Death before the modern age*. New York 1994, S. 28, 129.

49 Zit. nach Chalk/Jonassohn, 'The Conceptual Framework', in: Dies. (Hrsg.), *The History and Sociology of Genocide. Analyses and Case Studies*. New Haven, London 1990, S. 3-43, hier S. 12-32. Vgl. Schaller, *Genozidforschung*, (Anm.2), S. 15.

50 Barth, *Genozid*, (Anm.3), S. 27.

nach Fein Kriegstote, auch Guerrillakämpfer, als Opfer von Genoziden ausgeschlossen.⁵¹ Das Argument, dass Getätete zu „bewaffneten Rebellen“, „Partisanen“ oder ähnlichen Verbänden zählten, stellt jedoch eine häufig zu beobachtende Legitimationsstrategie genozidaler Staaten dar. Umgekehrt gilt: Wer sich wehrt, so die problematische Konsequenz der Feinschen Überlegungen, ist kein Genozid-Opfer.⁵²

Der Historiker *Frank Chalk* und der Soziologe *Kurt Jonassohn* betonen im Gegensatz zu Fein die Notwendigkeit einer staatlich organisierten oder ähnlich autoritären Macht:

„Genozid ist eine Form einseitiger Massentötung, mit welcher ein Staat oder eine andere Autorität versucht, eine Gruppe zu vernichten, nachdem diese Gruppe und die Mitgliedschaft in ihr durch den Täter definiert wurde.“⁵³

Diese Betonung des Staates beruht auf der Beobachtung, dass Genozide ein hohes Maß intensiver Planung und Organisation erfordern, die bislang nur von bürokratisch organisierten Mächten oder zumindest mit deren Rückendeckung verwirklicht werden konnten.⁵⁴ Ansonsten bringt auch dieser Definitionsneuersuch altbekannte Probleme: Ab welcher Größenordnung lässt sich von „Massentötung“ sprechen, und ist die Bestimmung von „Gruppen“ immer Bestandteil der Feindkonstruktionen auf Seiten der Täter? Für die Rechtsprechung wiederum plädiert Schabas dafür, eine positive Definition der „Gruppen“ aufrecht zu erhalten, dabei jedoch deren Konstruktionscharakter zu berücksichtigen. Damit wurden die vier in der Konvention genannten Kriterien (nationale, ethnische, rassische oder religiöse Kennzeichnung) „Eckpfeiler“⁵⁵ bilden, an denen sich die Anklage orientieren konnte, die aber nicht deterministisch zu verstehen seien.

Leider ist die dritte Fixierung der Konvention, die absichtsvolle Handlung seitens der Täter, bislang nicht auf das gleiche Problembewusstsein getroffen wie die anderen beiden Setzungen.⁵⁶ Juristisch hängt nämlich alles an diesem subjektiven Tatbestand.⁵⁷ Die erkennbare Absicht, eine Gruppe ganz oder teilweise zu zerstören, entscheidet darüber, ob das Verbrechen des Völkermordes vorliegt. Nun lässt sich nach juristischen Gepflogenheiten der Nachweis einer Absicht oder eines Vorsatzes nur auf natürliche Personen beziehen, nicht aber auf juristische (wie Staaten).⁵⁸ Im Extremfall, wenn keine individuelle Absicht nachweisbar ist, können millionenfache Morder nicht im Sinne der Konvention verurteilt werden. Und dieser Nachweis ist in aller Regel schwer zu führen, zum einen, weil Täter dazu neigen, alle Spuren zu verwischen und interne Entscheidungen sowie deren Strukturen

51 Helen Fein, *Definition and Discontent*, (Anm.59), S. 15. Vgl. auch Chalk, *Redefining Genocide*, (Anm.3), S. 49.

52 Eli; in Kurzat-Ahlers, *Ober das Töten in Genoziden*, (Anm.11), S. 192.

53 Frank Chalk/Kurt Jonassohn: *Genozid – ein historischer Überblick*. in: Mihran Dabab/Kristin Platt (Hrsg.): *Genozid und Moderne*. Bd. 1, *Strukturen kollektiver Gewalt im 20. Jahrhundert*. 410 S., Leske und Budrich, Opladen 1998, S. 300. Vgl. auch Horowitz, Irving Louis, *Taking Lives, Genocide and State Power*. New Brunswick 1980, S. 183-187.

54 Barth, *Genozid*, (Anm.3), S. 203. Vgl. Eli; in Kurzat-Ahlers, *Ober das Töten in Genoziden*, (Anm.11), S. 181-182. Weitere Definitionsvorschläge bei Israel W. Charny, *Toward a Generic Definition of Genocide*, in: George J. Andreopoulos (Hrsg.), *Genocide*. Philadelphia 1994, S. 64-94, hier S. 64; Verena Radkau, *Einleitung*, in: Dies. u. a. (Hrsg.), *Genozide*, (Anm.21), S. 14.

55 Schabas, *Genozid im Völkerrecht*, (Anm.6), S. 152. Andererseits verweist er, S. 709, darauf, dass eine restriktive Definition auch Vorteile gegenüber Missbrauch bote.

56 Chaumont, *Die Konkurrenz der Opfer*, (Anm.26), S. 177-179.

57 Schabas, *Genozid im Völkerrecht*, (Anm.6), S. 202ff.

58 Brumlik, *Theorie des Völkermords*, (Anm.40), S. 926. Vgl. auch Hans Vest, *Genozid durch organisierte Machtapparate. An der Grenze von individueller und kollektiver Verantwortlichkeit*. Baden-Baden 2002, S. 97.

zu verschleiern. Zum anderen ist massenhafte kollektive Gewalt ein sehr komplexes Phänomen, das sich aus Wechselspielen von Gelegenheiten und Bedingungen, von Initiativen in Zentrum und Peripherie speist.⁵⁹ Für den Holocaust war zwar Hitlers Wille entscheidend, aber auch dieser wurde schrittweise, an einzelnen Schauplätzen auf unterschiedliche Weise zu unterschiedlichen Zeiten realisiert, so dass man erst für die Zeit ab 1942 von einem Gesamtvorgang eines systematischen Massenmordes an den europäischen Juden sprechen kann.⁶⁰ Schließlich lässt sich anhand neuer Forschungen zur Gewalttheorie hinterfragen, ob der Begriff der Intention für Gewaltpraxen nicht einen geschlossenen Ablauf insinuiert, der Individualhandlungen möglicherweise gerecht wird, nicht aber komplexen Strukturen und diffus motivierten Strategien von kollektiven Massenmorden.⁶¹ Wissen die Täter tatsächlich zu jedem Zeitpunkt, was sie wollen oder wie sie dieses Ziel erreichen? Gibt es immer einen Moment der Entscheidung und der Beschlussfassung und wie entscheidend ist dieser für die Dynamik von Genoziden? Oder muss der Akt des Tötens, also die performative Ebene, stärker in die Analyse einbezogen werden?⁶² Diese performative Ebene stärker zu berücksichtigen, ist ein wichtiger Hinweis der Anthropologie. Wie in der konkreten Situation von Opfer und Täter verschiedenste Momente von Differenzproduktion, von Tötungslegitimationen (ob die von oben angebotenen oder die vor Ort konstruierten), von Ritualität, Expressivität und Inszenierung zusammentreffen, ist bislang zu wenig in den Blick genommen worden.

Noch in einer weiteren Hinsicht ist der Begriff der Intention fragwürdig: Was ist mit jenen kolonialen Kontexten, in denen Hunger, Krankheiten und Tod Ausdruck eben nicht von gezielter Planung, sondern von Planungsmängeln und organisatorischer Überforderung waren?⁶³ Gerade der Krieg gegen die Herero mit seinen verheerenden Folgen lässt sich nur unter Muhen als Genozid im Sinne der UN-Konvention definieren. Ob der deutsche General Lothar von Trotha von Beginn seines Amtsantritts an die Auslöschung der Herero anstrebte oder ob der Feldzug nicht eine militärisch aus dem Ruder gelaufene, auf frustrierter Hybris basierende Strafaktion war, ist umstritten.⁶⁴ Auch Helen Fein erinnert daran, dass viele Genozide sich erst im Verlauf längerer, teilweise widersprüchlicher und - *ex post* betrachtet - genozidaler Prozesse entfalten.

59 Die immer wieder erwähnte Dehumanisierung der Opfer ist übrigens auch kein Strukturmerkmal. Weder muss ein Völkermord immer mit einer Entmenslichung der „anderen“ einhergehen, noch führen Ideologien der Bestialisierung zwangsläufig zum Genozid. Vgl. Helen Fein, *Definition and Discontent. Labelling, Detecting, and Explaining Genocide in the Twentieth Century*, in: Stig Förster/Gerhard Hirschfeld (Hrsg.), *Genozid in der modernen Geschichte*. Münster 1999, S. 11-21, hier S. 14.

60 Steven T. Katz, der die intentionalistische Perspektive verfolgt, kann nicht nachweisen, dass es von Anfang an die Absicht der Nationalsozialisten war, alle Juden umzubringen. Vgl. Chalk/Jonassohn, *Genozid – ein historischer Überblick*, (Anm.29), S. 299.

61 Schaller, *Genozidforschung* (Anm.2); vgl. auch Gerlach, *Nationsbildung im Krieg*, (Anm.23), S. 351; Jacques Semelin, *Toward a vocabulary of massacre and genocide*, in: *Journal of Genocide Research* 5 (2003), S. 193-210, hier S. 198-200.

62 Alexander Hinton, in *Mittelweg* 36, 6 (2006); Semelin, *Vocabulary*, (Anm.61), S. 196.

63 Für einen engen Begriff: Schaller, *Genozidforschung*, (Anm.2), S. 14. Für einen weiten Begriff: David Maybury-Lewis, *Genocide against Indigenous People*, in: Hinton, *Annihilating Difference*, (Anm.29), S. 43-53.

64 Vgl. Barth, *Genozid*, (Anm.3), S. 128-136; Isabell V. Huli, *Military Culture and the Production of „Final Solutions“ in the Colonies: the Example of Wilhelminian Germany*, in: Gellately/Kiernan (Hrsg.), *The Specter of Genocide*, (Anm.2), S. 141-162. Gerade weil der Intensionsbegriff empirisch so wenig greifbar ist, überrascht Barths Fazit, als Charakteristikum genozidaler Gewalt des 20. Jahrhunderts die „unbedingte, gezielte und ideologische Vernichtungsabsicht“ in den Mittelpunkt zu stellen. Barth, *Genozid*, (Anm.3), S. 33.

2. Oie Shoah: Oer ultimative Genozid ader ein Völkermord in einer Kette von Völkermorden?

Wie sehr die Shoah die Geschichte des Begriffs Genozid beeinflusst hat, wurde schon erwähnt.⁶⁵ Nicht immer stieß diese Rezeption auf Zustimmung. Manche vermuten hinter der „Generalisierung“ des Holocausts in den Genozidtheorien Apologien und Relativierungen eines eigentlich unvergleichbaren „Zivilisationsbruchs“. Ein entschiedener Vertreter der Singularitätsthese ist *Yehuda Bauer*, der inzwischen dafür plädiert, den Terminus „Genozid“ für den Willen der Täter zu verwenden, einen Teil einer Gruppe auszurotten, während er den Begriff „Holocaust“ für den Willen, die gesamte Gruppe kompromisslos zu vernichten, reservieren möchte.⁶⁶ Obwohl diese Theorie bereits eine Korrektur ursprünglicher, vom politischen Israel noch immer vertretenen⁶⁷ Thesen darstellt, wonach einzig der Mord an den europäischen Juden den Terminus „Genozid“ für sich beanspruchen dürfe, wird Bauers Modell jedoch vorgehalten, dass es das grundsätzliche Problem nur auf eine andere Ebene verschiebe.⁶⁸ Jedes historische Ereignis ist einzigartig, gleichwohl spiegeln sich in ihm immer auch allgemeine Tendenzen, die zeitlich und räumlich zu entdecken sind. Aus dieser Einsicht wird die Theorie, wonach der Mord an den europäischen Juden unvergleichbar sei, inzwischen von der Mehrheit der Forschung abgelehnt. Sehr wohl nämlich lasse sich der Holocaust im Rahmen von schon längerfristigeren oder anderswo in Westeuropa auftauchenden Politiken der imperialen Ausweitung, von Plänen der Sozialtechnologie oder ideologischen Überzeugungen deuten. Zudem, so Boris Barth, erzeuge dieses Festhalten an der Beispiellosgkeit eine Sakralisierung oder Mystifizierung des „Holocaust“, die wissenschaftliche Zugänge eher verbaue.⁶⁹ Dementsprechend bezeichnet er den Konflikt um die Unvergleichbarkeit als auflösbar, indem die Synthese aus beiden Positionen gebildet werde. Die Shoah solte als ein Genozid unter anderen verstanden werden, der aber gleichwohl extreme Besonderheiten aufweise.⁷⁰ Die Höhe der Opferzahlen, die einmalige Art der Tötung und Verwertung von Opfern über ihren Tod hinaus, die umfangreiche Mobilisierung eines kompletten staatlich-bürokratischen Apparats zur Planung, Organisation und Durchführung des Massenmordes sowie die globale Totalität des Vernichtungswillens: alle diese Faktoren

65 Barth, Genozid, (Anm.3), S. 10, S. 52. Vgl. Gerhard Hirschfeld, Der Völkermord im 20. Jahrhundert. Plädoyer für eine vergleichende Betrachtung, in: Hartwig Hummel (Hrsg.), Völkermord. Friedenswissenschaftliche Annäherungen. Baden-Baden 2001, S. 78-90, hier S. 83.

66 Yehuda Bauer: Rethinking the Holocaust. 335 S., Yale University Press, New Haven/London 2001, S. 10f. Vgl. auch Steven T. Katz, The Holocaust in its historical context. Bd. 1, The Holocaust and Mass Death before the modern age. New York 1994; Dan Diner, Das Jahrhundert verstehen. Eine universalhistorische Deutung. München 1999, S. 195-249. Chaumont wiederum versucht mit seinem Modell, die Sonderstellung, die der Holocaust im Vergleich zu anderen Völkermorden einnimmt, durch den Zusammenfall von Ethnozid und Genozid zu erklären. Der Jüdöozid sei der bislang einzige Versuch, eine ethnozidäre Absicht mithilfe einer genozidären Praxis zu verfolgen. Chaumont, Die Konkurrenz der Opfer, (Anm.26), S. 185.

67 So entgegnete der damalige israelische Außenminister Shimon Perez in einem Interview mit einer türkischen Zeitung, „We reject all attempts to create a similarity between the Holocaust and the Armenian allegations. Nothing similar to the Holocaust occurred. It is a tragedy what the Armenians went through, but not a genocide.“ Vgl. Yair Auran, The banality of indifference. Zionism and the Armenian Genocide. New Brunswick, New Jersey 2002, S. 124; zit. nach Wojak/Meinl (Hrsg.), Völkermord und Kriegsverbrechen, (Anm.21), S. 13.

68 Zimmerer, Kolonialer Genozid?, (Anm.4), S. 111.

69 Vgl. Barth, Genozid, (Anm.3), S. 53f. mit weiteren Literaturangaben; Eric D. Weitz, Holocaust, Genozid und die Macht der Definition, in: Radkau u. a. (Hrsg.), Genozide, (Anm.21), S. 53.

70 Barth, Genozid, (Anm.3), S. 56.

stellen tatsächlich eine bislang ungekannte Dimension genozidaler Gewalt dar.⁷¹ Michael Geyer ergänzt zwei weitere Aspekte: Von anderen Fällen des Völkermords habe sich der Holocaust trotz aller Verwerfungen im Einzelnen durch die Vorsätzlichkeit und den Weitblick seiner Ausführung sowie die Dauer und Allgegenwart seiner Folgen unterschieden.⁷²

Wenn all diese Kennzeichen aber spezifisch sein sollen – und es auch sind –, dann fragt man sich, worin denn noch die Parallelen zu anderen Genoziden liegen. So einfach ist es mit der propagierten Synthetisierung offenbar nicht. Irritierenderweise bleibt auch nach etlichen vergleichenden Studien die Frage virulent, ob der Holocaust überhaupt *pars pro toto* für andere Genozide stehen kann. Die These, dass sich aus ihm allgemeine Entwicklungstendenzen ableiten ließen, ja dass er das Paradebeispiel für die Pathologien der Moderne sei⁷³, stieß bekanntlich auf einigen Widerstand. Verwiesen wurde – mit Blick auf andere europäische Nationen, die zwar auch „modern“ seien, derartige Gräueltaten aber nicht kennen – auf die deutschen Sonderbedingungen.⁷⁴ Und auch in der Genozidforschung drängt sich der Eindruck auf, dass die eminente Bedeutung, die dem Holocaust quasi als Idealtypus zugeschrieben wurde⁷⁵, den Blick auf die Analyse anderer Genozide eher verstellt. Nicht jeder Genozid ist – bei allem NS-eigenen Chaos – mit einer derartigen Präzision und Systematik als Ausrottungskampagne organisiert worden. Schon den Zeitgenossen fiel das auf. Die Washington Post schrieb am 3. Dezember 1944 mit Bezug auf Auschwitz: „Es ist vielleicht ein Fehler, diese Tötungen ‚Gräueltaten‘ zu nennen. [...] In dem von den Deutschen praktizierten Ausmaß handelt es sich um etwas Neues.“⁷⁶ Das grundlegende Problem bleibt damit bestehen: Viele Deutungskontexte der Shoah, etwa der Sozialdarwinismus oder die gesellschaftssanitären oder sozialtechnologischen Utopien, haben eine längere Tradition und ließen sich auch in anderen Nationen finden.⁷⁷ Aber weniger als diese Einzelmomente – das hat die neuere Forschung deutlich herausgearbeitet⁷⁸ – war es deren Verdichtung und mithin eine sehr spezifisch deutsche Entwicklung, die entscheidend für den Holocaust war: das die Gewalt forciierende Ineinandergreifen von Antisemitismus, Kriegführung, Besatzungspolitik, »Umvollungsplänen« (»Generalplan Ost«) und Ernährungspolitik.

71 Alan Rosenberg: Was the Holocaust unique? A peculiar Question? In: Isidor Walliman/Michael N. Dobkowski (Hrsg.): *Genocide and the Modern Age*. New York 1987, S. 145-161, hier S. 156; Robert F. Melson, Problems in the Comparison of the Armenian Genocide and the Holocaust. Definitions, Typologies, Theories, and Fallacies, in: Stig Färster/Gerhard Hirschfeld (Hrsg.), *Genozid in der modernen Geschichte*. Münster 1999, S. 22-35, hier S. 26. Eberhard Jäckel, Die elende Praxis der Untersteller. Das Einmalige der nationalsozialistischen Verbrechen läßt sich nicht leugnen, in: *Historikerstreit*. München 1987, S. 115-122, hier S. 118.

72 Konrad H. Jarausch/ Michael Geyer: *Zerbrochener Spiegel. Deutsche Geschichten im 20. Jahrhundert*. 493 S., Deutsche Verlagsanstalt, München 2005, S. 135. Vgl. auch Omer Bartov, *Mirrors of Destruction. War, Genocide, and Modern Identity*. Oxford 2000, S. 6.

73 Vgl. Zygmunt Bauman, *Dialektik der Ordnung. Die Moderne und der Holocaust*. Hamburg 1992.

74 Zuletzt Hans-Ulrich Wehler. Eine lebhaftige Kampfsituation. Ein Gespräch mit Manfred Hettling und Cornelius Torp, München 2006, S. 160-167.

75 Tatz, *With Intent to Destroy*, (Anm.21), S. 17, bezeichnet das nationalsozialistische Deutschland als „*Genocidal Engine*“. Vgl. auch Levene, *Jahrhundert der Genozide*, (Anm.29), S. 14.

76 Zit. nach Rabinbach, *Lemkins Schöpfung*, (Anm.15), S. 27.

77 Vgl. auch Ben Kiernan, *Twentieth-Century Genocides: Underlying Ideological Themes from Armenia to East Timor*, in: Gellately/Kiernan (Hrsg.), *The Specter of Genocide*, (Anm.21), S. 29-51.

78 Vgl. den guten Überblick bei Dieter Pohl, *Verfolgung und Massenmord in der NS-Zeit 1933-1945*, Darmstadt 2003; Ulrich Herbert (Hrsg.), *Nationalsozialistische Vernichtungspolitik 1939-1945. Neue Forschungen und Kontroversen*, Frankfurt/M. 1998.

II. Ergebnisse und Offene Fragen

Die Frage nach dem Stellenwert der Shoah ist eingebettet in eine ganze Reihe von Diskussionspunkten der vergleichenden Genozidforschung. Generell zeigt sich die Tendenz, das Phänomen Genozid immer weiter aufzufachern und in einzelne Aspekte zu zerlegen⁷⁹: Wie sehen die Abläufe von Völkermorden aus? Gibt es ein mechanisch wiederkehrendes Skript von Genoziden? Sind Genozide ein Phänomen der Moderne? Fordern Kriege die Dynamik von Völkermorden? Welche Rollen spielen Ideologien oder Religionen für die Täter? Was sind die Motive der Täter? Wie hoch sind die Opferzahlen? Inwiefern lasten auf den jeweiligen Gesellschaften Stressfaktoren wie Wirtschaftskrisen, Bevölkerungsexplosionen, Kriege im Nachbarland etc.? Wie haben die Gesellschaften versucht, nach dem Genozid mit dessen Folgen umzugehen?⁸⁰ Einige dieser Kontroversen werden im Folgenden vorgestellt.

1. Ein mechanisch wiederkehrendes Skript von Genoziden?

Der Vergleich von mehreren Fallstudien sowie der Wunsch, Ordnung in die konfuse Gewaltwelt des 20. Jahrhunderts zu bringen, haben zur Entwicklung mehrerer Verlaufsmodelle geführt. Sie sollen die notwendigen systematischen Schritte zur Vorbereitung von Genoziden erfassen. Gregory Stanton z. B. macht acht Etappen aus:

1. *Klassifikation* der Bevölkerung in verschiedene Gruppen
2. *Symbolisierung* von Unterschieden, z. B. durch kollektive Bezeichnungen („Juden“, „Zigeuner“ etc.) oder Kennzeichen (gelbe Sterne, besondere Kleidung etc.).
3. *Dehumanisierung* der intendierten Opfergruppe, teilweise gesetzliche Einschränkung der Menschenrechte für Angehörige der Gruppe, Verbreitung von Hass-Propaganda
4. *Organisation* durch Ausbildung spezieller Milizen, Einsatzgruppen, Aufbau der technischen Möglichkeiten zum Genozid
5. *Polarisation*, d. h. die Durchsetzung gesellschaftlicher Spaltung durch Terrormaßnahmen oder das Gegeneinander-Ausspielen verschiedener Gruppen
6. *Identifikation* der Opfer durch Todeslisten, Aussonderung, Konzentration von Angehörigen in gesonderte Wohnviertel oder Lager
7. *Extermination* der Opfergruppe, meistens durch Massenmord an den Mitgliedern der Gruppe
8. *Verleugnung* der Taten durch die Täter. Vernichtung von Beweisen, Flucht der Täter, oder Konstruktion von apologetischen Argumentationsmustern, die den Opfern die Schuld gibt.⁸¹

Am Beispiel des Genozids in Ruanda will er nachweisen, dass sein Modell zumindest auf ein Verbrechen anwendbar sei, das er als „the clearest case of genocide since the Ho-

79 Vgl. unveröffentlichtes Paper von Andreas Mehler zum Thema „Vergleich“, Hamburger Institut für Sozialforschung, 12. April 2005; das Konzept der Konferenz Verbrechen gegen die Menschheit. Ursachen, Formen und Prävention von Völkermord, am 3.-5.11.2005, Berlin, <http://boell.de>.

80 Vgl. zu diesem Punkt z. B. die Beiträge von Beatriz Manz sowie May Ebiara und Judy Ledgerwood über Guatemala bzw. Kambodscha in Hinton, *Annihilating Difference*, (Anm.29).

81 Gregory Stanton: Wie wir Genozid verhindern können. Der Aufbau einer internationalen Kampagne zur Beendigung von Genoziden. in: Verena Radkau u. a. (Hrsg.): *Genozide und staatliche Gewaltverbrechen im 20. Jahrhundert*. 176 S., Studien-Verlag, Wien u. a. 2004, S. 29-39, hier S. 31. Auf die Verleugnung als wichtige Phase eines Genozids weist auch Colin Tatz hin; vgl. ders., *With Intent to Destroy*, (Anm.21), S. 122ff.

locust" bezeichnet.⁸² Gert Hankel erteilt diesem Phasenmodell jedoch eine Absage, weil es Eindeutigkeit und Zwangsläufigkeit suggeriere, die es so nicht gebe. Die Dynamik von Massenmorden speise sich häufig weniger aus quasi folgerichtigen Radikalisierungsstufen, sondern aus anderen teils kontingenten, teils chaotischen, teils nichtintendierten Beschleunigungsfaktoren wie externen Eskalationsschüben oder kriminellen Bereicherungsstreben.⁸³ Diese Skepsis in Bezug auf allzu schematische Verlaufsmodelle und die Betonung externer Einflüsse teilt Hankel mit Semelin. Der französische Soziologe unterstreicht zudem an den Beispielen Deutschland, Ruanda und Bosnien die Passivität der internationalen Gemeinschaft als weiteren Eskalations-Faktor.⁸⁴

2. Genozid - ein modernes Phänomen?

Genozid, so eine vielfach vertretene These, sei ein altes Phänomen, das einen neuen Namen trage.⁸⁵ Norbert Finzsch beispielsweise differenziert zwischen vormodernen und modernen Genoziden.⁸⁶ William O. Rubinstein wiederum schlägt ein flinfstufiges Periodisierungsmodell vor, das auch vormoderne Genozide einschließen soll:

1. Genozide in vorschriftlichen Gesellschaften
2. Genozide im Zeitalter der Imperien und Religionen (500 v. Chr. bis 1492)
3. Koloniale Genozide (1492-1914)
4. Genozide im Zeitalter des Totalitarismus (1914-1979)
5. zeitgenössische ethnische Säuberungen und Genozide (1945 bis heute).⁸⁷

Aber wie bei jedem Begriff, der erfunden ist, stellt sich die Frage, wie zweckmäßig es ist, ihn, der mit Blick auf spezifische Phänomene des 20. Jahrhunderts entworfen wurde, auf frühere Zeiten zurückzuprojizieren – und damit zu dekontextualisieren.⁸⁸ Die Thesen Arno Meyers, die strukturelle Ähnlichkeiten zwischen dem nationalsozialistischen Vernichtungskrieg, mittelalterlichen Kreuzzügen und dem Dreißigjährigen Krieg postulieren, erscheinen Barth daher als abwegig.⁸⁹ Auch andere Autoren sprechen sich dafür aus, dass „Genozid“ nur Phänomene des 20. Jahrhunderts beschreiben kann, das damals „Jahr-

82 Gregory H. Stanton, Could the Rwandan Genocide Have Been Prevented? In: Schaller/Boyidjian u. a. (Hrsg.), Enteignet. Vertrieben. Ermordet, (Anm.2), S. 437-456, hier S. 437 u. S. 439-443.

83 Vgl. das unveröffentlichte Paper von Gerd Hankel zur Konferenz Verbrechen gegen die Menschheit. Ursachen, Formen und Prävention von Völkermord, am 3.-5.11.2005, Berlin, <<http://boell.de>> [Zugriff am 10.02.2007].

84 Semelin, Purifier, (Anm.21), S. 150-168.

85 David Maybury-Lewis, Genocide against Indigenous Peoples, in: Alexander L. Hinton (Hrsg.), Annihilating Difference, (Anm.29), S. 43-53, hier S. 43. Maybury-Lewis bezieht z. B. die Vernichtung Karthagos durch das Römische Reich ein.

86 Norbert Finzsch, Genocides against Native Americans between Individualist Agenda and State-Implemented Program, in: Stig Förster/Gerhard Hirschfeld (Hrsg.), Genozid in der modernen Geschichte. Münster 1999, S. 48-59, hier S. 48-50. Vgl. auch David E. Stannard, American Holocaust. The Conquest of the New World. Oxford 1992, S. 256.

87 William D. Rubinstein, Genocide, (Anm.21), S. 6. Vgl. Auch Chalk/Jonassohn, Genozid – ein historischer Überblick, (Anm.29), S. 305-308.

88 Vgl. Barth, Genozid, (Anm.3), S. 35.

89 Arno Meyer, Krieg als Kreuzzug. Das Deutsche Reich, Hitlers Wehrmacht und die „Endlösung“. Reinbek 1989. Ähnlich auch Sven Lindqvist, Durch das Herz der Finsternis. Ein Afrika-Reisender auf den Spuren des europäischen Völkermords. Frankfurt am Main 1999. Vgl. Barth, Genozid, (Anm.3), S. 36f. Zur Kritik an Meyers These siehe auch Christopher Browning, The Path to Genocide. Essays on Launching the Final Solution. Cambridge 1992.

hundert der Genozide" etikettiert wird.⁹⁰ Will die Genozidforschung jedoch der vermuteten qualitativen und quantitativen Ausweitung des Phänomens Massenmord im 20. Jahrhundert gerecht werden, so muss sie erklären, welche historischen Rahmenbedingungen gerade das „Jahrhundert der Extreme" zu einem Zeitalter der Gewalt machten und worin die moderne Spezifik dieser Massentötungen im Vergleich zu früheren Ereignissen liegen soll.

Hierzu sind im Wesentlichen zwei gegenläufige Thesen publiziert worden. Raphael Lemkin ging davon aus, dass Genozide eine Regression zu vergangener und längst überwunden geglaubter Barbarei darstellen.⁹¹ Sie seien somit ein altes Phänomen, das gleichwohl in der Moderne eine neue Qualität gewonnen habe. Demgegenüber hält Zygmunt Bauman, angeregt durch die „Dialektik der Aufklärung", ⁹² den Genozid für ein genuin modernes Produkt der instrumentellen Vernunft. In der Regel werde er in einem „rationalen" Akt geplant und von einer umfassend organisierten, durchgeplanten, wissenschaftlich orientierten und technisch effizienten Mordmaschinerie durchgeführt.⁹³ Darüber hinaus sei ein grundlegender Unterschied zwischen emotional begründeter Heterophobie und theoretisch-pseudowissenschaftlich legitimiertem (und damit genuin modernem) Rassismus festzustellen. Vernichtungspolitik im 20. Jahrhundert zeichnet sich gegenüber älteren Vorläufern – etwa den Ausrottungsfeldzügen der Assyrier, den Massakern der Kreuzritter oder der Conquista Lateinamerikas – dadurch aus, dass sie einem zweckrationalen Plan folge und auf dem höchsten technischen Niveau durchgeführt werde. Zudem ermögliche eine bürokratische Organisation mit moderner Arbeitsteilung den „Schreibtischtätern", ihre Aufgaben ohne Berücksichtigung ethischer Vorbehalte durchzuführen.⁹⁴ Diese Sichtweise orientiert sich fast ausschließlich an einem „industrialisierten" Typus von Tat. Nun lässt sich aber gerade für den Holocaust das Neben- und Miteinander von vielen Mordmethoden feststellen, etwa das bürokratisch sinnlose Oualen von Menschen oder Massen- wie EinzelerchieBungen. Moderne Ratio ist nur ein Aspekt.

Entscheidender ist, dass sich das Verständnis von Kriegführung und damit von erlaubter, gebotener und verbotener Gewalt im 19. und 20. Jahrhundert gewandelt hat, nicht zuletzt aufgrund der negativen Erfahrung von Zerstörung und Massenmord. Erst auf dieser Folie eines Deutungswandels im Wertekomplex heutiger Gesellschaften, nämlich der Entwicklung der Menschenrechte und des Glaubens an eine universale Menschenwürde, etablierte sich der Begriff „Genozid". Damit spiegelt die Konjunktur des Terminus auch die enorm gestiegene mediale Präsenz (entgrenzter) kriegerischer Gewalt. Mit Radio, Femse-

90 Eric Weitz, *A Century of Genocide* (Anm.21); Michael Mann, *The Dark Side of Democracy. Explaining Ethnic Cleansing*, Cambridge 2005; Radkau u. a. (Hrsg.), *Genozid*, (Anm.21); wenig konsistent: Levene, *Jahrhundert der Genozide?*, (Anm.29). Vgl. auch Schaller, *Genozidforschung*, (Anm.2), S. 9; Eric Hobsbawm, *Das Zeitalter der Extreme. Weltgeschichte des 20. Jahrhunderts*, München, Wien 1995, S. 26.

91 Raphael Lemkin, *Genocide – a modern crime*, in: *Free World. A Non-Partisan Magazine devoted to the United Nations and Democracy* 4 (1945), S. 39-43. Vgl. Barth, *Genozid*, (Anm.3), S. 33; Levene, *Jahrhundert der Genozide?*, (Anm.29), S. 9.

92 Max Horkheimer/Theodor W. Adorno, *Dialektik der Aufklärung. Philosophische Fragmente*. Amsterdam 1947.

93 Bauman, *Dialektik der Ordnung*, (Anm.73); ders., *Modernity and the Holocaust*, in: Hinton (Hrsg.), *Genocide. An anthropological reader*. Malden, Mass. 2002, S. 110-133. Vgl. auch Micha Brumlik, *Das Jahrhundert der Extreme*, in: Fritz Sauer Institut (Hrsg.), *Völkermord und Kriegsverbrechen in der ersten Hälfte des 20. Jahrhunderts*. Frankfurt am Main 2004, S. 19-36, hier S. 26; Peter Imbusch, *Moderne und Gewalt. Zivilisationstheoretische Perspektiven auf das 20. Jahrhundert*, Wiesbaden 2005.

94 Vgl. auch Vest, *Genozide durch organisatorische Machtapparate*, (Anm.58), S. 30.

hen, Internet ist es möglich, vorausgesetzt man will es, sich über jede Massaker in der Welt zu informieren. Mit diesen Bilderfluten wird das Synthetisierungsbegehren, die Informationen zu einer kohärenten Geschichte der Völkermorde zusammenzubinden, drängender als je zuvor. Der Terminus entsprang mithin gleichermaßen dem Ordnungswunsch wie dem Wunsch nach Einhegung von Gewalt. Alle Täter werden in diesen Werte- und Normenhorizont gestellt. Dieser Referenzrahmen lässt sich aber nun für Akteure der Antike, des Mittelalters oder der Frühen Neuzeit nicht herstellen. Gewalt kannte damals andere Grenzen, es gab damit andere Gewaltexzesse, die innerhalb der damaligen Gewalt- und Normensysteme zu erklären sind. Insofern kann der Terminus Genozid nur moderne Phänomene reflektieren, weil er in einem modernen Spannungsfeld generiert wurde und agiert: nämlich dem Spannungsfeld von Menschenrechten und deren Verletzung. Ein „anachronistischer“ Zugriff auf „Genozide“ in vormodernen Zeiten könnte dennoch durchaus interessant sein. Er wäre nur eine methodische Entscheidung, die begründet werden muss. Sein Zweck ist dann jedoch nicht mehr, „Genozide“ zu erklären, sondern große Gewaltereignisse.

Ob nun das 20. Jahrhundert tatsächlich auch mehr Massenmorde gesehen hat, und wenn ja, was hierfür ursächlich sein könnte, ist eine zweite Frage. Die Erkenntnisse der bisherigen Genozidstudien legen nahe, dass das 20. Jahrhundert ein sehr spezifisches Setting an Strukturen bereit stellte, das eher eine Ballung von entgrenzter Gewalt ermöglichte als andere Jahrhunderte zuvor: z. B. eine bürokratische Rationalität, eine effektivere Militärtechnologie, massenmediale Propagandamöglichkeiten, neue Ideologien, der Expansionsdrang und Imperialismus des Westens. Zu klären wäre dann, unter welchen Bedingungen manche Akteure einzelne oder alle Bestandteile dieses Settings aufnehmen und Minderheiten zu störenden und bedrohlichen „Anderen“ erklärten, gegen die „etwas“ unternommen werden musste. Genauso wichtig wäre aufzuschlüsseln, wann diese Prozesse abgebrochen wurden, wann es also zu einer Umkehr oder einem Innehalten der Täter kam. Wie wird die Geschichte der Gewalt zu einer „Geschichte der nie ausgeübten Gewalt“⁹⁵?

3. Der Staat als Täter

Dass erst die Moderne die Mittel für Genozide zur Verfügung gestellt hat, meint auch der Historiker Peter Fritzsche. Diese Modernität bedeute aber auch, dass Genozide in der Regel Staatsverbrechen seien:

„In the end [...] genocide is not the sum of many atrocities. It rests on substantial intellectual work to revisualize the population, to dramatize national history as both something mortally imperilled and potentially transformed, purified, and existentially sanctioned [...], and to overrule pattern of neighbourliness. It is also something extralocal, requiring the power and resources of the state in order to keep the killings from coming to an end and to keep the killers from going home. Revolutionary vision, ideological commitment, and state power are needed to sustain genocidal practices across large territories over many months and even years. It is unlikely that this sort of equation was possible before the nineteenth century.“⁹⁶

Genozide als komplexe, z.T. über Jahre währende Prozesse setzen in der Regel intensive Planungs- und Organisationsphasen voraus, die aller Erfahrung nach nur Staaten zur

95 Sven Hillenkamp, *Die Bombe Mensch*, in: *Die Zeit* vom 11.05.2006. Vgl. auch Midlarsky, *The Killing Trap*, (Anm.21). Er benachteiligt Fälle, in denen es *nicht zum Genozid* kam.

96 Peter Fritzsche, *Genocide and Global Discourse*, in: *German History* 23 (2005), Nr. 1, S. 96-111, hier S. 105.

Verfügung stellen können. Aus dieser Beobachtung wurde schon früh der Schluss gezogen, dass insbesondere totalitäre Regime genozidanfällig seien.⁹⁷ Wenn man sich aber die wesentlichen Völkermorde des 20. Jahrhunderts anschaut, wird man neben Staaten wie NS-Deutschland oder die UdSSR mit dem Osmanischen Reich, dem Irak, Kambodscha, Indonesien, Ruanda und Burundi auch keineswegs totalitäre, sogar nicht immer autoritäre Regime finden. Ebenso war keineswegs immer ein psychopathischer Diktator ausschlaggebend.⁹⁸ Und ob Genozide vor allem ein Krisenphänomen von sich unter Druck sehenden starken Staaten sind, wie *Manus Midlarsky* und *Jacques Semelin* vermuten⁹⁹, bedarf ebenfalls weiterer Forschungen. Im Hinblick auf die NS-Führung und ihren Entschluss zur Ermordung der europäischen Juden zumindest boten sich auch andere Sichtweisen an. Das Argument, dass mittlerweile belegt sei, dass die NS-Führung im Dezember 1941 angesichts der drohenden Niederlage gegen die Sowjetunion die „Endlösung“ initialisierte, unterschlägt zweierlei: Zum einen gibt es Gegenansichten, die den Entschluss auf Herbst, also auf den Höhepunkt deutscher Erfolge und damit einer Euphorie des „alles ist möglich“ datieren. Zum anderen verkennt die Rede vom Krisenphänomen die Dynamik der Shoah. Die sowjetischen Juden z. B. wurden schon seit Beginn des Unternemens Barbarossa im Juni 1941 ermordet – anfanglich die Männer, dann unterschiedslos alle, einschließlich Frauen und Kinder.

Gegen die These vom „starken Staat als *hostis populi*“ konnten die Beispiele der *failed states* geltend gemacht werden.¹⁰⁰ Lassen wir zunächst die Debatte darum, ob diese Staaten überhaupt je als Staaten agierten bzw. ob sie nicht eine sehr eigene Staatlichkeit kennzeichneten, außer Acht. Hier geht es darum, dass gerade nicht *strong states* agieren, die Liber ein gut ausgebautes Gewaltmonopol verfügen, es aber missbrauchen, sondern dieses ist nur höchst unvollständig verwirklicht. Daraus rühre, so die These, eine Privatisierung der Gewalt, deren unheilvolle Dimensionen noch nicht abzuschätzen sind. Bislang ist es aber noch zu keinem Genozid in einem „failed state“ gekommen.¹⁰¹ Angesichts dieser Ergebnisse kann man der Folgerung zustimmen, dass ein stabiler Rechtsstaat mit demokratischer Verfassung Genozide kaum denkbar macht.¹⁰² Mit *Klaus Schlichte* wäre aber darauf zu beharren, den Staatsbegriff zu dynamisieren und in dem scheinbar unitären Akteur das Machtfeld deutlich herauszuarbeiten, in dem verschiedene Individuen und Institutionen um Geltung streiten.¹⁰³

97 Vgl. Helen Fein, Genozid als Staatsverbrechen. Beispiele aus Ruanda und Bosnien, in: Zeitschrift für Genozidforschung 1 (1998), Nr. 1, S. 36-45, hier 37-38.

98 Levene, Jahrhundert der Genozide, (Anm.29), S. 25.

99 Manus I. Midlarsky: The Killing Trap. Genocide in the Twentieth Century. 463 S., Cambridge University Press, Cambridge 2005, betont, dass die Täter oft geglaubt hätten, aus einer Position der Schwäche heraus handeln zu müssen. In Bedrohungslagen neigten insbesondere Staaten, die einer machiavellistischen realpolitischen Doktrin folgen, dazu, bereits geschehene oder befürchtete Verluste auszugleichen, indem sie imaginäre oder reale Feinde prophylaktisch eliminieren. Je stärker die politischen Unsicherheitsfaktoren für ein Regime seien, desto wahrscheinlicher werde ein Genozid. Ähnlich Jacques Semelin: Toward a vocabulary of massacre and genocide, in: Journal of Genocide Research 5 (2003), S. 194.

100 Martin Gilbert, Twentieth-Century Genocide, in: Jay Winter (Hrsg.), America and the Armenian Genocide of 1915, Cambridge 2003, S. 9-36.

101 Vgl. zum Fall Darfur bzw. Sudan: Gerard Prunier, Darfur. Der »uneindeutige« Genozid, Hamburg 2006.

102 Vgl. Barth, Genozid, (Anm.3), S. 175-176.

103 Klaus Schlichte: Der Staat in der Weltgesellschaft. Politische Herrschaft in Asien, Afrika und Lateinamerika. 329 S., Campus, Frankfurt/M. 2005.

Aber keine These in der Genozidforschung, die nicht ohne Widerspruch bliebe. Die Immunität von Demokratien hat *Michael Mann* kürzlich in Zweifel gezogen. Er betont, dass die Repräsentationsidee des liberalen westlichen demokratischen Staates in Mittel- und Osteuropa häufig im ethnonationalen Sinn umgedeutet wurde. Das demokratische Konzept sei zugunsten eines fragmentierten Nationenprinzips aufgegeben worden, das Teile der Bevölkerung vom eigentlichen „Volk“ ausschloss. Ethnische Säuberungen und Genozide seien somit die „Schattenseite der Demokratie“.¹⁰⁴ Nun lässt sich an dem Band vieles kritisieren, insbesondere, dass er außer den USA als Fallbeispiele keine Demokratien behandelt. Schwerer wiegt, dass Mann eigentlich ethnische Säuberungen erklären mochte, d. h. wie moderne Staaten das Problem ethnischer Vielfalt kreieren bzw. sich ihm stellen. Der Oxford-Soziologe verliert sich aber in der Dichte des historischen Materials. Vertreibungen, Genozide, Massaker, es geht wie bei so manch anderem Autor trotz vorgeblich genauer Systematik alles bunt durcheinander. Samantha Power und *Adam Jones* sind in dem Punkt Demokratien und exzessive Gewalt zumindest etwas konziser. Sie wenden ein, dass Demokratien des westlichen Typus, in erster Linie die Vereinigten Staaten von Amerika, sich häufig nicht nur passiv verhalten, wenn außerhalb ihres Einflussbereichs Genozide verübt werden,¹⁰⁵ sondern oftmals verbrecherische Regime in ihren (zumindest genozidverdächtigen) Aktionen unterstützen, wenn sie nicht sogar selbst genozidale Gewalt anwenden.¹⁰⁶ Aus der Regierungs- oder Gesellschaftsform eines Staates lasse sich somit keine zuverlässige Prognose über dessen „Genozidanfälligkeit“ ableiten. Dies zeige auch die Gegenprobe, denn selbst in totalitären Diktaturen stelle Genozid die Ausnahme dar.¹⁰⁷ Es ist jedoch ein Faktum, dass in der zweiten Hälfte des 20. Jahrhunderts stabile Demokratien in der Nachbarschaft reifer Demokratien keine Massenmorde an ihrer Bevölkerung begangen haben. Quintessenz all dieser Befunde wäre demnach, dass eine labile, undemokratische Staatsform einer der Faktoren sein konnte, der zu Genoziden beiträgt.¹⁰⁸

4. Kriege und Genozide

Obwohl eine wesentliche Neuerung der Konvention darin bestand, Völkermord nicht mehr an zwischenstaatliche kriegerische Handlungen zu binden, bildet der Zusammenhang zwischen Genozid und Krieg einen weiteren Fokus der Genozidforschung. Es ist umstritten, ob bestimmte kriegerische Akte unter Umständen als genozidal eingestuft werden sollten und ob Kriege Genozide begünstigen.¹⁰⁹ So stellt der Professor für Internationale Beziehungen an der University of Sussex, *Martin Shaw* fest, dass kriegerische Handlungen die Tendenz aufweisen, zu „degenerieren“, und spricht in diesen Fällen von „genozidalen Kriegen“.¹¹⁰ Auch der Bombenkrieg der Alliierten als absichtsvolle Tötung deutscher und japanischer

104 Michael Mann: *The Dark Side of Democracy. Explaining Ethnic Cleansing*. 590 S., Cambridge University Press, Cambridge 2005, S. 1-33, S. 68f. Vgl. auch Christine Schweitzer/Björn Aust/Peter Schlotter (Hrsg.), *Demokratien im Krieg*, Baden-Baden 2005.

105 Power, *Problem from Hell*, (Anm.13).

106 Adam Jones (Hrsg.): *Völkermord, Kriegsverbrechen und der Westen*. 534 S., Parthas, Berlin 2005. Der Sammelband führt zahlreiche Kriegsverbrechen, Verbrechen gegen die Menschlichkeit und Genozide des 20. Jahrhunderts auf, an denen „der Westen“ direkt beteiligt war.

107 Barth, *Genozid*, (Anm.3), S. 175; Valentino, *Final Solutions*, (Anm.21), S. 28.

108 Vgl. auch Wolfgang Merkel, *Im Zweifel für den Krieg*, in: *Die Zeit* vom 20.04.2006.

109 Vgl. Barth, *Genozid*, (Anm.3), S. 37-41.

110 Martin Shaw: *War and Genocide. Organized Killing in Modern Society*. 272 S., Polity Press, Cambridge 2003.

Nichtkombattanten im Zweiten Weltkrieg steht nach Meinung des Soziologen Leo Kuper und anderer in diesem Kontext.¹¹¹ David E. Stannard und Ward Churchill haben dafür plädiert, neben Hungersnot, Seuchen auch Besatzungsherrschaften in den Begriffskatalog aufzunehmen.¹¹² Die anthropologische Genozidforschung schließlich weist darauf hin, dass die kulturelle Konstruktion von Differenzen sowohl im Krieg als auch im Genozid nach ähnlichen Gesetzmäßigkeiten verläuft.¹¹³ Die Frage sollte aber nicht sein, ob einzelne Strategien genozidal waren, sondern ob die Kriegführung etwa der USA im Zweiten Weltkrieg – aber auch in Vietnam – danach trachtete, alle Deutschen oder alle Japaner umzubringen – aber ob nicht selbst diese Kriegsverbrechen immer noch dem Ziel verhaftet waren, den Krieg zu beenden und Frieden zu schließen.¹¹⁴ In der Logik von Genoziden kommen Friedensschlüsse nicht vor, weil auf der gegnerischen Seite keiner mehr da sein sollte, mit dem man diesen Akt begehen konnte. Gleichwohl können die Grenzen zwischen der Totalisierung von Kriegen und dem Aufkommen genozidaler Gewalt fließend sein¹¹⁵, zumal wenn, wie im 20. Jahrhundert, Kriege immer mehr nicht mehr nur gegen Armeen, sondern gegen Zivilisten geführt werden.¹¹⁶ Deswegen ist zu vermuten, dass Kriege mit ihrer militärischen, organisatorischen, informationstechnischen und propagandistischen Aufrüstung stets das Potenzial zur Entgrenzung von Gewalt aufweisen.¹¹⁷ Anders ausgedrückt: Krieg schafft einen neuen Möglichkeitsraum. Jacques Semelin betont hier zwei Aspekte: Zum einen seien Kriege wichtig für eine Radikalisierung der Feindbilder, ob innere oder/und äußere Feinde. Zum anderen werde unter dem Dogma der „nationalen Sicherheit“ nun vieles möglich, was zuvor nur auf Unverständnis gestoßen sei.¹¹⁸ Semelins Fingerzeig auf die Bedeutung von Sicherheit- bzw. Unsicherheitsgefühlen ist brisant, zeigt er doch einmal mehr, dass weniger „sachrationale“ Begründungen als vielmehr Gefühle bzw. Visionen von wiederzugewinnender Sicherheit auf Seiten der Täter handlungsleitend sein können. Zu analysieren bliebe aber dennoch, unter welchen Umständen Täter diese Optionen aufnehmen bzw. vor allem, wann sie es nicht tun.¹¹⁹

5. Ideologien als Antriebsmomente - oder als Dynamisierungsfaktoren?

Auch Eric D. Weitz' Ziel ist es, einerseits zu erklären, inwieweit gerade das 20. Jahrhundert als „Jahrhundert der Genozide“ gelten kann, und andererseits Typologien zu erstellen, mit deren Hilfe Kategorien für eine vergleichende Genozidforschung bereitgestellt werden konnten.¹²⁰ In vier Länderstudien pointiert er die modernen Ideologien „Rasse“ und „Nation“ als *conditio sine qua non* für Genozide: Das sowjetische Imperium strebte eine klassenlose Groß-Nation an,¹²¹ das nationalsozialistische Deutschland eine „rassisch“ homogene Volks-

111 Leo Kuper, *Genocide. Its political use in the twentieth century*. New Haven 1981, S. 17, S. 46.

112 Nach Rabinbach, *Lemkins Schöpfung*, (Anm.15), S. 23.

113 In der Regel durch die Konstruktion binärer Gegensatzpaare, vgl. Hinton, *Dark Side*, (Anm.2g), S. 8ff.

114 Vgl. auch Yves Ternon, *Der verbrecherische Staat. Völkermord im 20. Jahrhundert*. Hamburg 1995, S. 63.

115 Barth, *Genozid*, (Anm.3), S. 180.

116 Monks, *Die neuen Kriege*, (Anm.10), S. 91ff.

117 Barth, *Genozid*, (Anm.3), S. 181f.

118 Semelin, *Purifier*, (Anm.21), S. 168.

119 Helen Fein, *Genocide. A Sociological Perspective*. London 1993, S. 94ff.

120 Eric O. Weitz: *A Century of Genocide. Utopias of Race and Nation*. 360 S., Princeton University Press, Princeton 2003.

121 Ebd., S. 54-63.

gemeinschaft,¹²² die Roten Khmer versuchten, die vermeintlich glorreiche Vergangenheit eines mittelalterlichen Reiches wiederherzustellen und verknüpften dieses Streben mit einer kommunistischen Ideologie,¹²³ während Serbiens Regierung von einem ethnisch homogenen serbischen Großreich träumte.¹²⁴ Weitz. macht für jeden der genannten Genozide eine Phase aus, in der geklärt werden musste, wer überhaupt zur entsprechenden Feindgruppe gehörte.¹²⁵ Obwohl die Ideologien in der Regel „objektive“ Kriterien vorgaben – der Nationalsozialismus die biologische Rasse, der Kommunismus die Akkumulation von Produktionsmitteln – weist Weitz. nach, dass häufig soziale Kriterien (politische Überzeugung der Opfer, Umstände des Lebenswandels etc.) bei der Auswahl der „Feinde“ eine mindestens ebenso große Rolle spielten. Problematisch aus Sicht der Täter war hierbei, dass sich eine vielfältig heterogene Bevölkerung zumeist nicht in einfache Raster einteilen ließ, was in der Praxis zu widersprüchlichen Einteilungen führte.¹²⁶

Die nächsten Phasen nennt Weitz. „Purging the Population“ und „The Ultimate Purge“,¹²⁷ wobei sich die Bezeichnung auf eine langsam eskalierende Selektions- und Diskriminierungspraxis bezieht. Am Beispiel Jugoslawiens ist Weitz. freilich gezwungen, seine Terminologie zu verlassen, indem er das entsprechende Kapitel „Preparing for Population Purges“ nennt. Im Unterschied zu den anderen vorgestellten Systemen, die versuchten, ihre Existenz über Krisen hinwegzureretten, spielte in Jugoslawien die Transformation des Systems selbst eine zentrale Rolle für die Ausweitung der Gewaltmaßnahmen. Gemeinsam sei, so Weitz. resumierend, dass jedes Völkermordende Regime von einer revolutionären Utopie besessen war. Zur Verwirklichung dieser Vision brauchte es den modernen Staat, vor allem aber die Partizipation großer Teile der Bevölkerung.

Der Historiker *Omer Bartov* führt diese Erklärungsansätze noch weiter. Er untersucht, inwiefern die Konstruktion von Identität in der Moderne kriegs- und genozidale Gewalt beeinflusst hat. Unter den Rahmenbedingungen einer kulturell bedingten Verherrlichung kompromissloser Gewalt sei es ein Kernbestand europäischer Identität des 20. Jahrhunderts gewesen, einer Bedrohung des eigenen Wesens durch die gezielte Vernichtung eines realen oder imaginierten Feindes zu begegnen, der gleichzeitig als Widerpart des eigenen Daseins diene.¹²⁸ Darüber hinaus stelle Gewalt ein zielgerichtetes Mittel dar, nicht nur die eigene, sondern auch fremde Identität zu definieren. Dies könne bis zur Aberkennung einer überhaupt menschlichen Identität reichen. Ein „*project of remaking humanity*“,¹²⁹ eine „Bio-

122 Ebd., S. 103-114.

123 Ebd., S. 145-159.

124 Ebd., S. 191-201.

125 Ebd., S. 63-68, S. 114-119, S. 159-164, S. 201-205.

126 So zeigen die Definitionskriterien der Nürnberger Gesetze, dass es nicht nur von der Abstammung, sondern auch von der religiösen Praxis bzw. sozialen Faktoren (etwa Eheschließungen) abhing, ob ein Mensch als „Volljude“, „Halbjude“ oder „Vierteljude“ gezählt wurde. Ebd., S. 116.

127 Ebd., S. 68-74, S. 119-124, S. 164-170.

128 Omer Bartov: *Mirrors of Destruction. War, Genocide, and Modern Identity*. 302 S., Oxford University Press, Oxford 2000, S. 91-142. Vgl. auch Wilshire, *Get 'em all, kill 'em!*, (Anm.21), S. 105. Wilshires nicht gerade bescheidener Anspruch besteht darin, eine Genozid-Theorie mit zuverlässiger Vorhersagekraft zu entwickeln. Seine wenig kontextualisierende Studie unterscheidet nicht mehr zwischen Genozid und Terrorismus und erklärt Völkermorde aus objektiv vorhandener kultureller Differenz. Diese Thesen vermögen nicht zu überzeugen. In den meisten multikulturellen Gesellschaften kommt es gerade nicht zu Genoziden. Das nationalsozialistische Deutschland z. B. war vor 1939 in ethnischer Hinsicht recht homogen, das Regime konnte nur unter hohem logistischen Aufwand feststellen, wen es denn eigentlich umbringen wollte.

129 Bartov, *Mirrors of Destruction*, (Anm.128), S. 5f.

politik" großen Ausmaßes, konnte maßgeblich zu einem Genozid beitragen.¹³⁰ Gegen Bartov wird eingewandt, dass er trotz aller phänomenologischen Aufmerksamkeit und kulturtheoretischen Präzision die Erkenntnisse der Totalitarismustheorie nicht überschreitet, und dass er sich ebenso wie diese den Vorwurf gefallen lassen müsse, dass Ideen alleine nicht soziale und politische Praxen erklären.¹³¹

Nun kann man mit *Kieser* und *Schaller* zusammenfassen, dass sich seit dem Ende des 19. Jahrhunderts wichtige Wahrnehmungsmuster entwickelten, die das Denken der Zeitgenossen empfänglich für das Konzept „Völkermord“ machten. Sozialdarwinismus, Nationalismus und Rassismus begünstigten die Übertragung biologischer Metaphern sowie medizinischer, bakteriologischer und chirurgischer Bilder auf menschliche Identität und zwischenmenschliches Zusammenleben.¹³² Trotzdem bleiben bei manchen Studien Zweifel. Ist dem Faktor Ideologie tatsächlich immer der hohe Stellenwert einzuräumen, den nicht zuletzt eine mitunter eher an älterer Fachliteratur denn an Quellen geführte Rezeption der Shoah-Forschung vorgibt?¹³³ Oder wie Peter Fritzsche formulierte: „And still, while genocide is racism, racism is not genocide: racial categories have to be mobilized for genocide to occur; they are necessary but not sufficient components of an explanation for mass murder.“¹³⁴ Der rassistische Süden der Vereinigten Staaten im 19. und frühen 20. Jahrhundert hat ebenso keinen Völkermord an seiner schwarzen Bevölkerung versucht wie Südafrika nach 1948.¹³⁵ Auch leiden viele der Genozidstudien unter ihren uneingestanden Prämissen, z. B. dass zwischen Wahrnehmungen und Handeln nicht unendlich viele Brüche sein können, oder dass rassistische Utopien immer auf ihre blutige Umsetzung drängten. Vielleicht sind die Kausalitäten mitunter umgekehrt: Erst im Zuge der Gewaltausübung entwickelt sich aus einem rassistischen Weltbild Rassenhass¹³⁶, oder erst im Nachhinein werden Legitimationen für gewaltvolles Handeln gesucht.

6. Die Täter

Diese Kritik aufnehmend wendet sich die anthropologische und psychologische Forschung zunehmend von einer objektiv-äußeren Strukturbeschreibung des Genozids ab und den kulturellen Konstruktionen von Differenz zu. Ein Augenmerk gilt dabei der Symbolik von

130 Vgl. Brumlik, *Jahrhundert der Extreme*, (Anm.93), S. 25.

131 Brumlik, *Theorie des Völkermords*, (Anm.40), S. 927.

132 Hans-Lukas Kieser/Dominik J. Schaller: *Völkermord im historischen Raum 1895-1945*. in: dies. (Hrsg.): *Der Völkermord an den Armeniern und die Shoah*. 648 S., Chronos, Zürich 2002, S. 11-80, insbes. S. 11, S. 59.

133 Vgl. Barth, *Genozid*, (Anm.3), S. 172ff.

134 Fritzsche, *Genocide and Global Discourse*, (Anm.96), hier S. 97.

135 Barth, *Genozid*, (Anm.3), S. 183.

136 Isabel V. Hull, *The Military Campaign in German Southwest Africa, 1904-1907*, in: *GHI Bulletin* 37 (2005), S. 39-44, hier S. 42. Auch George Steinmetz, *Von der Eingeborenenpolitik zur Vernichtungsstrategie: Deutsch-Südwestafrika, 1904*, in: *Peripherie* 97/98 (2005), S. 195-227, zweifelt für den Herero-Krieg an der Deutungskraft des Rassismus-Modells. Er macht darauf aufmerksam, dass nur selten danach gefragt wird, wie fest etabliert diese Bezugssysteme waren und was mit konkurrierenden Modellen passierte. In Windhoek gab es, bis von Trotha kam, mit Leutwein einen Mann an der Spitze, der zwar auch in rassistischen Kategorien dachte, aber letztlich auf Ausgleich zwischen den Herero und den Deutschen bedacht war. Steinmetz sieht die Ursache für die Radikalisierung daher eher in der Psyche von Trothas bzw. in Spannungen zwischen Mittelklassen und Aristokratie im Reich, die auf die Kolonie übertragen wurden.

Gewalt, der jeweiligen Bedeutung von Körpern und deren Zerstörung.¹³⁷ Einen weiteren Schwerpunkt bilden die mentalen Dispositionen der Täter und ihrer Motivation.¹³⁸

Anregend sind in diesem Zusammenhang die Überlegungen von Jacques Semelin. Er betont, dass die diskursive Unterfütterung von Völkermorden häufig von dem Begriff der „Reinheit“ dominiert wird, der sowohl in einer rassistischen, nationalen, politischen oder religiösen Konnotation erscheine. Die Vernichtung der so gebrandmarkten „Anderen“ erscheint gemäß dieser Logik dann nicht nur als Präventions- und Überlebensmaßnahme, sondern werde im Zuge von Identitätskonstruktionen zum konstitutiven Beweis der „eigenen“ Identität.¹³⁹

Tatsächlich jedoch könnte in einer Konkretisierung des sozialen Aktes Töten, der Handlung wie des Sprechens über kollektive Gewaltverbrechen, wie sie etwa *Thomas Kuhne* und *Peter Gleichmann* einfordern, ein Schlüssel liegen, die verwirrenden und zähen Definitionsdebatten über den Begriff „Genozid“ aufzubrechen.¹⁴⁰ Was bringt Menschen dazu, anderen Menschen so viel Leid zuzufügen, andere Körper so zu zerstören? Welche Handlungsräume sehen die Täter? Welche Netzwerke oder Komplizengemeinschaften von Gewalt bilden sie? Sind diese gruppenspezifischen und sozialpsychologischen Prozesse von Kameradschaftserzeugung Vorbedingung für das gemeinsame Morden? Wird das persönliche Moralempfinden durch eine „Gemeinschaftsmoral des Normbruchs“ ersetzt?¹⁴¹ Welche Zusammenhänge gibt es zwischen Gewalt und Lust, sexueller wie nichtsexueller? Welche körperlichen Erfahrungen machen Täter beim Töten?¹⁴² Alle diese Fragen sprechen Themen an, die nicht nur in Genoziden von Bedeutung sind, sondern generell für eine Theorie der Gewalt. Daneben könnte für die Forschung von Relevanz sein, stärker als die Intention der Täter das Verhältnis oder das Ineinandergreifen von persönlichen Motiven wie etwa Habgier, Lust oder gekränkter Eitelkeit, und nationalistischen, rassistischen usw. Überzeugungen und Legitimationen zu problematisieren. Wichtig wäre zudem, nach den verschiedenen Tätertypen, die für entgrenzte Gewaltakte verantwortlich sind, zu differenzieren: Der Entscheidungsträger, der (vielleicht) nicht tötet, der Propagandist, der ebenfalls nicht direkt tötet, der Organisator und schließlich die Exekutoren, die „tatnachen“ Täter.¹⁴³ Schließlich sollten die Verlaufsformen und Dynamiken einzelner Gewaltextzesse, etwa in Massakern, untersucht werden. Wann agieren Einzelne, wann Gruppen, wie viele beteiligen sich, wer

137 Vgl. etwa Christopher C. Taylor, *The cultural face of terror in the Rwandan genocide of 1994*, in: Hinton, *Annihilating Difference*, (Anm.2g), S. 137-178.

138 Harald Welzer, *Täter. Wie aus ganz normalen Menschen Massenmörder werden*. Frankfurt a. M. 2005; Christopher R. Browning, *Ganz normale Männer, das Reserve-Polizeibataillon 101 und die „Endlösung“ in Polen*, Reinbek bei Hamburg 2005. Orig.-Ausg., New York 1998.

139 Semelin, *Vocabulary*, (Anm.61), S. 197.

140 Peter Gleichmann/Thomas Kuhne (Hrsg.): *Massenhaftes Töten. Kriege und Genozide im 20. Jahrhundert*. 418 S., Klartext, Essen 2004, insbesondere die Einleitung: Thomas Kuhne, *Massen-Töten. Diskurse und Praktiken der kriegerischen und genozidalen Gewalt im 20. Jahrhundert*, S. 11-52. Vgl. auch Hans Rudolf Schelling, *Genozid, Sozialpsychologische Erklärungsansätze*, in: Schaller/Boyadjian u. a. (Hrsg.), *Enteignet. Vertrieben. Ermordet*, (Anm.2), S. 2g-66; Alain Berta-lo, *Van der Ethnisierung zum Genozid. Mechanismen der Mobilisierung Unbeteiligter zu Akteuren kollektiver Gewaltextzesse*, in: Schaller/Boyadjian u. a. (Hrsg.), *Enteignet. Vertrieben. Ermordet*, (Anm.2), S. 67-74; Semelin, *Vocabulary*, (Anm.61), S. 19g,

141 Thomas Kuhne, *Massen-Töten*, (Anm.140), S. 37.

142 Alf Ludtke, *War Work. Aspects of Soldiering in Twentieth Century Wars*, in: ders./ Bernd Weisbrod (Hrsg.), *No Man's Land of Violence. Extreme Wars in the 20th Century*. Göttingen 2006, S. 127-151.

143 Semelin, *Vocabulary*, (Anm.61), S. 203.

wird warum zum *bystander*, welche Akte werden ausgeübt, wann und wie kommt es zu einem Inne- oder Einhalten der Gewalt?¹⁴⁴

Die genannten Ansätze versprechen einerseits eine schlussigere Einordnung des Themas „Genozid“ in den Gesamtzusammenhang der Gewalt- und Konfliktforschung des 20. Jahrhunderts.¹⁴⁵ Andererseits ist in ihnen eine Spannung zwischen Mikro- und Makroperspektive zu beobachten. So werden einer zu individualisierten Perspektive ihre Anfälligkeit für subjektive Darstellungen und Entkontextualisierung vorgeworfen.¹⁴⁶ Der Kriminalsoziologe *Alex Alvarez* versucht, diese Klippen zu umschiffen, indem er sowohl die Verantwortlichkeit der Regierenden als auch individual- und kollektivpsychologische Verdrängungsprozesse gegenüber Leid und Unrecht berücksichtigt.¹⁴⁷ Dieses Vorgehen ist in seiner Umsicht nur zu unterstreichen. Die subjektiven Deutungen und Erfahrungen sollten an das Gesamtsetting der Gewalt zurückgebunden werden. Tätermotivation und situative Dynamiken ohne die politischen Rahmenbedingungen, ohne Entscheidungsstrukturen und Entschlussbildungen wären nur die halbe Geschichte.

Resümee

Die vergleichende Genozidforschung hat unser Wissen über Genese, Strukturen, Verläufe, beteiligte Institutionen, Motivationen und Legitimationen von Völkermorden erheblich erweitert. Paradoxerweise hat dieses Wissen aber kaum zu einer Präzisierung der Theorie von Genoziden geführt. Das hat vor allem zwei Ursachen. Erstens hat der Genozid-Begriff mehrere Stoßrichtungen. Er soll sowohl empirische Befunde erklären und zu einer Phänomenologie von Völkermorden beitragen als auch normativ wirken. Kaum ein Forschungsbegriff ist somit so vorbelastet wie dieser Terminus. Zweitens gehört zur Eigentümlichkeit exzessiver Gewalt im 20. Jahrhundert offenbar eine außerordentliche Heterogenität des Phänomens. Beide Momente zusammen haben dazu geführt, dass Genozid zu einem generellen Begriff für Massenverbrechen geworden ist. Der Wunsch war, eine knappe Formel für höchst komplexe Phänomene zu haben. Immerhin räumt selbst Leo Kuper, einer der Begründer der vergleichenden Genozidstudien, ein, es gebe „keinen einheitlichen genozidalen Vorgang, und damit keine Basis für eine allgemeine Theorie des Genozids“¹⁴⁸. Die Frage ist, ob der Begriff des Genozids dann die Trennschärfe liefern kann, die sich die Wissenschaft von ihm erhofft. Es mag Fälle mit einigen Gemeinsamkeiten geben, etwa der Mord an den Armeniern und den Juden, aber rechtfertigen diese wenigen Parallelen die Typologisierung durch einen einzigen Begriff? Wird die Forschung je über das abgestufte Vokabular verfügen, um die unterschiedlichen Fälle historiographisch genau einzuordnen und der jeweiligen Wahrnehmung der Betroffenen gerecht zu werden? Taugt der überdehnte Begriff

144 Vgl. hierzu demnächst Bernd Greiner, *Krieg ohne Fronten. Die USA in Vietnam*, Kapitel 7: 16. März 1968, Hamburg 2007.

145 Vgl. auch Alf Ludtke/Bernd Weisbrod (Hrsg.), *No Man's Land of Violence. Extreme Wars in the 20th Century*. Göttingen 2006.

146 Dan Diner, *Die Wahl der Perspektive. Bedarf es einer besonderen Historik des Nationalsozialismus?* In Wolfgang Schneider (Hrsg.), *Vernichtungspolitik*. Hamburg 1991, S. 65-75, hier S. 72.

147 Alex Alvarez: *Governments, Citizens, and Genocide. A Comparative and Interdisciplinary Approach*. 240 S., Indiana University Press, Bloomington 2001.

148 Leo Kuper, *Genozid*, zit. nach Levene, *Jahrhundert der Genozide*, (Anm.2g), S. 14.

noch zu Erklärungen oder ist Alain Finkielkraut zuzustimmen, der den Terminus unter die Rubrik „verbale Inkontinenz“¹⁴⁹ einordnet?

Orei Aspekte sprechen fUr ein Aufgeben des Begriffes oder zumindest fUr eine Losung von seinen fesselnden Schwächen und damit fUr mehr empirische und theoretische Offenheit:

1. Nimmt man die Genozidkonvention als Grundlage, dann ergeben sich vielfältige Probleme. Die Konvention fasst sehr heterogene Massentötungen zusammen, lässt andere aber außen vor. Diese schwer vermittelbare Logik irritiert vor allem in Hinblick auf das daraus entstehende Allerlei kollektiven Gewaltüberschusses, dazu im dritten Punkt mehr. Noch schwerer wiegt die juristisch zwar notwendige, empirisch aber schwierig nachzuweisende Setzung, dass die Täter stets vorsätzlich gehandelt haben. In Armenien oder in Deutsch-Südwestafrika entsprang massenhafter Tod vielleicht eher Indifferenz und Unfähigkeit. Im Ergebnis mag es gleich sein, ob Millionen systematisch umgebracht werden oder ob ihr Tod billigend in Kauf genommen wird. Für die Forschung, die sich um Ursachen und Gründe bemüht, macht es aber einen Unterschied. So erinnern Kriminologen daran, dass die Böse-verursacht-Böses-Gleichung nur allzu oft eine Täuschung ist. Ein zweiter Aspekt: Zumindest partiell sind entgrenzte Gewalttaten im 20. Jahrhundert dadurch charakterisiert, dass sie durch kein Kalkül mehr gedeckt waren. Massenhafte Tötung scheint zwar immer einer instrumentellen Logik zu folgen, die „Feinde“ müssen zum eigenen Heil endgültig verschwinden. Die Dynamik der Shoah, aber auch anderer Gewaltereignisse wie Massaker legt aber nahe, dass Gewalt sich auch jenseits dieser Instrumentalität bewegt, dass sie Selbstzweck wurde.¹⁵⁰

2. Die alte Frage, ob Massenmorde, die sich zu verschiedenen Zeiten in unterschiedlichen Kulturen und unter unterschiedlichen Bedingungen abgespielt haben, ein universales Merkmal der Menschheitsgeschichte sind, kann vermutlich mit Ja beantwortet werden. Aber diese Erkenntnis bleibt banal. Erklärt werden soll doch, warum es 1993 zum Völkermord in Ruanda kam – warum nicht schon früher und warum nicht in Nigeria? Warum kam es zu diesen Gräueln in Ost-Timor, warum zu diesem Terrorregime in Kambodscha? Warum fiel Jugoslawien Anfang der 90er Jahre aus dem europäischen Rahmen der Friedfertigkeit? Sicher lassen sich immer wieder bestimmte Einzelfaktoren destillieren, die in vielen Fällen eine Rolle gespielt haben: Krisen, Kriege, Rassismus, ethnische Neuordnungspläne, totalitäre Systeme, Diktatoren usw. Aber viele Beispiele lehren, dass nicht diese Einzelfaktoren als solche, ja nicht einmal deren Zusammentreffen Genozid auslösend sind, sondern dass die Dynamik der Gewalt der zeitlichen wie räumlichen Verdichtung dieser Faktoren, ihrer wechselseitigen Beeinflussung, entspringt. Denkt man diese Erkenntnis weiter, kann es keine Gesetzmäßigkeiten geben, nach denen Genozide funktionieren. Es scheint, als ob allein die Aussage, dass stabile Demokratien in einem stabilen demokratischen Umfeld wenig genozidanfällig seien, einige Solidität besitzt. Weitere Wahrscheinlichkeitsaussagen bleiben das, was sie sind: Gänge auf dünnem Eis.¹⁵¹ Umgekehrt gilt damit aber auch: Der

149 Alain Finkielkraut, *The Future of Negation. Reflections on the Question of Genocide*, Lincoln 1998, S. 95.

150 Vgl. Ferdinand Sutterluty, Ist Gewalt rational?, in: *WestEnd. Neue Zeitschrift für Sozialforschung*, Jg. 1, Heft 1/2004, S. 101-115.

151 Barths Diktum, Genozid, (Anm.3), S. 183: „Rassismus ist eine notwendige Voraussetzung für Genozid“, ist angesichts seiner differenzierten Argumentation irritierend. Ob z. B. in Ruanda *rassistische* Vorurteile eine Rolle spielten, ist ja umstritten.

Gang der Dinge ist zu beeinflussen. Hätte sich die Weltgemeinschaft mehr in Ruanda engagiert, wäre der Völkermord zumindest in dieser Form nicht passiert.

3. Der wissenschaftliche Impetus der Genozidforschung besteht zu einem Gutteil darin, die Gewaltgeschichte der letzten beiden Jahrhunderte systematisch und differenziert zu erfassen und dabei vor allem deren Exzesscharakter auf die Spur zu kommen. Aber gerade die Vielschichtigkeit der Phänomenologie von entgrenzter Gewalt erschwert unser Verstehen und setzt Klassifizierungen enge Grenzen. Exzessive Gewalt mag sich darin ähneln, dass eine Gruppe von Menschen zum „Problem“ wird, das die tonangebende Gruppe „loswerden“ will. Schon bei Terror allerdings greift diese Definition nicht. Es geht nicht um die Existenz der Gegnergruppe *per se*, sondern um deren Handeln. Vermeintlich subversiven Akten soll vorgebeugt werden. Warum also Menschen zu Gegnern erklärt werden, mit welchen Merkmalen sie als Feind ausgestattet werden, welches die Prinzipien der Selektion sind, all dies unterscheidet sich bei näherem Hinsehen. Historisch betrachtet war exzessive Kollektivgewalt ein höchst variables Ereignis, selbst wenn das Ergebnis immer gleich entsetzlich anmuten mag. Insofern ist Gewalt, wie Welzer betont, sozial und historisch spezifisch, quantitativ wie qualitativ.¹⁵² Ihre Einzelphänomene, ob sie nun als ethnische Säuberungen, Massaker, Terror, Deportationen, Pogrome, counter-guerilla oder Völkermord im engeren Sinne der Auslöschung willkürlich definierter Gruppen von Menschen¹⁵³ bezeichnet werden, wiesen nicht nur jeweils sehr spezifische, sich voneinander unterscheidende Logiken auf, sie waren auch nicht immer miteinander verknüpft. Empirisch lassen sich nur wenige Fallstudien finden, in denen „Vergewaltigung, Folter und Mord [...] am Ende der Skala [standen], die mit Ausgrenzung, Stigmatisierung, Beraubung und Entrechtung“¹⁵⁴ begann. Genauso wenig lässt sich eine Rangordnung von Massenverbrechen bilden. Welche Kriterien sollten hierfür auch den Maßstab bilden? Die Zahl der Toten, die Art der Tötung? Gegenwärtig erscheint vermutlich vielen weniger Völkermord als islamisch-fundamentalistischer Terrorismus als das ultimativ Böse.¹⁵⁵

Letztlich könnten diese Zweifel zum Eingeständnis führen, dass Strukturen, Täter, Opfergruppen von Fall zu Fall so unterschiedlich gelagert waren, dass es kaum oder sogar keine Gemeinsamkeiten von Genoziden – nicht einmal im 20. Jahrhundert – gibt. Weitere Definitionsversuche von Genoziden jenseits der UN-Konvention erscheinen jedenfalls wenig forschungsdienlich. Wenn es aber starke empirische Gründe gibt, den Begriff in der Forschung fallen zu lassen, was macht man dann mit der normativen Seite des Begriffes? Was wurde sich ändern, wenn dieser oder jener Mord kein Genozid gewesen wäre? Juristisch gibt es mit den Straftatbeständen Verbrechen gegen die Menschheit und dem Kriegsstrafrecht weiterhin Handhabe gegen die Täter. Politisch könnte es größere Probleme geben.

Nehmen wir das Beispiel Armenien. Nach wie vor ist umstritten, ob für den Völkermord ein groß angelegter, sukzessiver Vernichtungsplan oder Unfähigkeit und Überforderung der osmanischen Behörden ursächlich war.¹⁵⁶ Lassen wir uns einmal für einen Augenblick darauf ein, dass es kein Völkermord im Sinne der Konvention gewesen ist. Auf den ersten

152 Welzer, Täter, (Anm.138), S. 258.

153 Vgl. Barth, Genozid, (Anm.3), S. 7.

154 Benz, Vermeidung, (Anm.29).

155 Vgl. Semelin, Vocabulary, (Anm.61), S. 193.

156 Vgl. Donald Bloxham, The Great Game of Genocide. Imperialism, Nationalism, and the Destruction of the Ottoman Armenians, Oxford 2005; Guenter Lewy, The Armenian Massacres in Ottoman Turkey. A Disputed Genocide, Salt Lake City 2005; ferner Hans-Lukas Kieser, Urkatastrophe am Bosphorus. Der Armeniervölkermord im ersten Weltkrieg als Dauerthema internationaler (Zeit-)Geschichte, in: Neue politische Literatur 50 (2005), Nr. 1, S. 217-234.

Blick wären die politischen Konsequenzen enorm. Zu statuieren, dass an den Armeniern nicht das absolute Verbrechen begangen wurde, wurde alle Apologeten in Hurraschreie ausbrechen lassen. Aber die Schlussfolgerung, dass dies vielleicht kein Genozid im Sinne der Konvention war, impliziert nicht, wie der Jurist Jom Axel Kämmerer jungst betonte, dass das den Armeniern bereitete Schicksal volkerrechtskonform war.¹⁵⁷ Tod durch Unterlassung ist nicht akzeptabler als Mord. Lapidare Repliken und Schonreden der Türkei konnten weiterhin unter Verweis auf das damalige internationale wie nationale Recht und den immer noch wirksamen Rechtsfolgen ausgehebelt werden. Insofern konnte die juristische, politische wie wissenschaftliche Argumentation durchaus subtiler verlaufen, ohne dass man dieses Staatsverbrechen des osmanischen Reiches relativiert oder gar auf die Ausfluchte und Verdrehungen der Regierung in Ankara einschwenkt.¹⁵⁸

Wissenschaftlich scheint die Tragfähigkeit des Genozidbegriffes erschöpft, ironischerweise nicht zuletzt, weil seine Durchsetzung die Aufmerksamkeit auf das Phänomen massiver Gewalt gelenkt und sich unser Kenntnisstand enorm verbreitert hat. Eben weil wir jetzt so viel mehr wissen, legt er der Forschung Fesseln an. Mit seinen nur scheinbar klaren Vorgaben verstellt er den Blick auf die mitunter doch sehr anders gelagerten Realitäten entgrenzter Gewalt. Lässt man ihn für den wissenschaftlichen Diskurs fallen, konnte man sich endlich zu der Einsicht durchringen, dass Gewaltabläufe *auch* inkonsistent und kontingent, dass die Handlungen der Opfer, des Auslands eine Rolle spielen können für Entscheidungsprozesse der Täter, und dass nicht immer der Wille entscheidend ist, sondern die Tat.

Anschrift der Verfasserin: PD Dr. Birthe Kundrus, Hamburger Institut für Sozialforschung, Mittelweg 36, 20148 Hamburg.

Email: Birthe.Kundrus@his-online.de

Anschrift des Verfassers: Henning Strotbek, Neutorstraße 78, 26721 Emden.

Email: henning.strotbek@gmx.de

157 Jörn Axel Kämmerer, Was geschah in Armenien, in: FAZ 24.04.2006, S. 42.

158 Rabinbach, Lemkins Schöpfung, (Anm.15), S. 24, weist im Übrigen darauf hin, dass die Verpflichtung zur Intervention Staaten davon abhalte, den Begriff öffentlich zu gebrauchen. Insofern sei er vielleicht eher handlungslähmend als -leitend.

Merkzettel

Minimalpunkte für die Anfertigung einer Arbeitsdisposition für Magister- und Doktorarbeiten

Das Geld der Stiftungen wird immer knapper; unsere Zeit als Prüfer wird von immer mehr Examenskandidaten nachgefragt. Um die Erfolgsaussichten einer Stipendienbewerbung verlässlicher abzuschätzen – und/oder um einen Überblick über Ihr Arbeitsprojekt gewinnen zu können, der uns erlaubt, bei möglichen Problemen gleich zu Anfang korrigierend eingreifen zu können – sollten Sie der Anfrage um die Anfertigung eines Stipendiengutachtens oder der Vergabe eines Abschlußthemas folgende Informationen beifügen:

- 1) Provisorische Formulierung des gewünschten/vorgeschlagenen Themas
- 2) Problembeschreibung
- 3) Forschungsstand (knappe Auseinandersetzung mit der einschlägigen Literatur)
- 4) Erkenntnisinteresse
- 5) Fragestellung und Arbeitshypothesen
- 6) Vorgeschlagener Gang der Untersuchung
- 7) Quellen, empirische Erhebungen usw.
- 8) Methoden (der Auswertung des Materials wie der Anfertigung der Darstellung)
- 9) Mögliche Ergebnisse
- 10) Voraussichtliche Gliederung
- 11) Zeitplan für die Arbeit

Vollständige Postanschrift, email-Adresse und Telefon nicht vergessen, damit wir Sie bei Rückfragen erreichen können !

Sicherlich wird die Anfertigung eines Exposés nach diesem Muster zu Anfang etwas mehr Zeit kosten; während des Prozesses des Verfassens der Arbeit selbst kann es Ihnen aber immer wieder von Nutzen sein. Und: Sie sparen sich und uns eine Menge Arbeit, wenn Sie sich an diese Punkte halten: unvollständige Dispositionen geben wir nämlich in Zukunft postwendend zurück.

Fragen zum Aufbau von Arbeiten und zur Zitierweise beantworten im übrigen verschiedene Veröffentlichungen des Instituts, die Sie aus dem Netz downloaden können unter:

<http://egora.uni-muenster.de/pol/service/sic/download.shtml>

In der Auswahl Ihrer Zitierverfahren sind Sie im Rahmen des Gesagten, das Sie in den auf der Institutswebsite angebotenen Broschüren finden können, frei: nur *einheitlich* sollte Ihr Verfahren schon sein.

R.M.



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

School of Political Science and International Studies

CRICOS Provider Number 00025B

ESSAY GUIDE 2017



POLITICAL SCIENCE AND INTERNATIONAL STUDIES

For more information please contact:

School of Political Science and
International Studies

Email: polsis@uq.edu.au

Telephone: + 61 7 3365 2858

Website: www.polsis.uq.edu.au

Contents

Section 1: How to Construct an Essay.....	4
Planning	4
1. Choosing an Essay Topic.....	4
2. Understanding the task.....	4
3. Making an Argument	5
4. Outlining or Planning Your Essay.....	5
5. Reading	6
6. Taking Notes	7
7. Writing and Revising	7
Writing Conventions.....	8
Word count	10
Presentation	10
Section 2: Referencing	11
Which referencing system should you use?	11
Style 1: Author-date in-text referencing system.....	11
Simple in-text citations with an identified author.....	11
More complicated instances of in-text citation	11
Reference List/Bibliography.....	12
Book.....	12
Multi-authored book	12
Edited Book.....	13
Chapter in an Edited Book.....	13
Scholarly Journal Article	13
Newspaper and magazine where the author is named, and the source is hard copy	13
Newspaper and magazine where the author is not named and the source is hard copy.....	13
Online source that is only available online, not in hard copy.....	13
Government Report with named author	14
Government Report with institutional author	14
United Nations Document.....	14
Translated Work	14
Reprints of older Work	14
Book Review	14
Unpublished Work.....	14
Style 2: Footnote referencing system.....	15
Examples of footnotes.....	15
Books with one, two or three authors	15
Books with four or more authors	15
Edited book.....	15
Chapter in an Edited Book.....	15
Scholarly journal article.....	16
Scholarly journal article that is available online with a doi	16
Newspaper and magazine article where the author is named, and the source is hard copy	16
Newspaper and magazine article where the author is not named and the source is hard copy.....	16
Newspaper and magazine article that you accessed online.....	16
Material sourced online that is only available online	16
Government Report with named author	16
Government Report with institutional author	16
United Nations Document.....	16
Translated book.....	16
Book review	17
Unpublished Work.....	17

<i>Re-citing items that have already appeared in an earlier footnote.....</i>	<i>17</i>
<i>Reference List/Bibliography.....</i>	<i>17</i>

SECTION 1: HOW TO CONSTRUCT AN ESSAY

This Essay Guide is designed to help you plan and construct a standard essay in the School of Political Science and International Studies. You may be asked to do other assessment pieces by your course coordinator that might require different types of preparation than the ones outlined here. In all cases, follow the advice of your course coordinator as to the exact type of assessment that is being requested of you.

Planning

Before you start writing your essay it is really important that you take the time to plan your essay.

There are seven major steps to planning good essays:

1. Choosing an Essay Topic
2. Understanding the Task
3. Making an Argument
4. Outlining or Planning the Essay
5. Reading
6. Taking Notes
7. Writing and Revising

1. Choosing an Essay Topic

You should **choose a topic early** in the semester and begin working on it with the intention of producing more than one draft. A common mistake is to prepare essays in a frantic, last-minute rush. Effective **time management is essential** to successful tertiary study. Therefore, you must consider the work required in all your courses and plan accordingly, especially when you have multiple essay deadlines that fall around the same time.

2. Understanding the task

The most common problem in undergraduate essays is not doing precisely what the assignment asks you to do. Please pay close attention to the assessment as described in the Electronic Course Profile.

Once you understand what your task is you should break it down into its component parts. This enables you to decide what material is relevant. Suppose, for example, the following question was asked: 'Is a two-party system necessary for the existence of representative democracy?' A careful analysis of the question might suggest that a suitable answer could focus on components such as:

- the nature of representative democracy
- the role of political parties in representative democracy
- features of representative democracy strengthened by the role of political parties
- the idea that representative democracy is neither dependent on, nor weakened by, a two-party system; and a conclusion setting out your evaluation of these points.

Some of the components of an essay topic may not come to mind immediately. When thinking about a topic refer to lecture notes, your course textbook or related reading material to get a better idea of the

topic(s) at hand. However, do not cite lectures in your assignment and make sure you read more widely than the required textbook.

3. Making an Argument

One of the **most common mistakes made when writing an essay is failure to make a main argument**. A main argument is a response to the essay question combined with a justification for your answer. This should guide how the essay is structured.

In order to explain why you are making your argument in a particular way, you need to demonstrate that you have researched the topic. This requires use of academic sources. You are expected to demonstrate understanding of the issue or subject by drawing upon ideas, theories, research findings and related information that supports your argument. This **requires an analytical, not a descriptive approach**, so it is not sufficient to simply reproduce relevant information or repeat other people's arguments. Nor is it appropriate to answer an essay question with broad generalisations (for example 'democracy is desirable') for which no supporting evidence or reasoning is provided. Remember that a scholarly argument or claim cannot merely be asserted. It must be substantiated by **evidence** and supported by **authoritative sources**. It is also important to acknowledge alternative viewpoints.

Be critical in your approach to the topic: In courses dealing with politics you are expected to be critical in the sense of determining whether or not the evidence available justifies the conclusions that are drawn from it; in courses dealing with political ideas, you are expected to question the assumptions involved in the material. Being critical also implies identifying gaps in others' arguments or the evidence they use and postulating alternative explanations or interpretations. Being critical might also involve examining the limitations of your own views or addressing counterarguments that arise in relation to the argument presented in your essay.

4. Outlining or Planning Your Essay

Once you have analysed the question, you should organize the ideas into an outline. The outline should ensure that the essay has a logical structure. It also facilitates the preparation of the essay by guiding your reading, note taking and writing. There is more than one way to write an outline. One is to do a diagrammatic/brain map approach, putting boxes on the page for each component (eg introduction, one paragraph per key point you want to make in the essay, conclusion), and adding dot points for each point you want to make in each component. Another is to do a linear, narrative plan in which you use headings for the introduction, each paragraph, and the conclusion and put dot points under each heading for what you plan to write in that component.

It is important at this stage to keep in mind the **stipulated word limit** (see the section on Word Count on page 8 of this Guide) and any other requirements set by the Course Coordinator. It may help to assign an approximate number of words for each section of your outline. By recognising the points that are central and those that are peripheral to your argument and by allocating appropriate word-lengths, your outline will provide a useful guide for how much reading and writing are required.

Be prepared to **revise your argument and your essay outline**. In the course of reading widely it may become evident that your initial argument or outline is incomplete or inadequate in some way. Revise it as new information and perspectives come to your attention.

5. Reading

A good essay requires wide reading. Reading widely will help you develop the breadth of knowledge necessary to evaluate ideas and arguments put by others. However, mere quantity is not enough; you should choose your material intelligently and be selective.

When you start your reading, it will be most useful to get a general idea about what the book or the article is about to determine whether you want to draw on it in your essay or not. Therefore, when you are starting out, you might want to read the introduction and conclusion of a chapter or an article to get an overall sense of its approach. You can also use a book's index, or the sub-headings within chapters, to identify relevant information. As you narrow down your topic, you can then focus more on the details and on the paragraphs, quotations or ideas most relevant to your argument.

It is usually best to read from the general to the specific. Begin by reading the relevant sections of introductory texts, and then move on to more detailed publications or specialised journal articles. Individual courses and assignments may require different amounts of reading. Note that it is normally impossible to write a convincing essay based upon a limited number of sources.

Drawing on Academic Sources: Most Course Coordinators issue reading lists to help students choose relevant material. These are the **best place to start** researching an essay topic. Further references may be compiled by using bibliographies in books and journal articles.

A good quality essay makes effective use of academic sources, such as books, chapters in edited volumes and peer reviewed journal articles. Textbooks are useful as they enable you to provide a broader context or to illustrate the relevance of the essay question. However, they typically provide only short references to academic debates and literatures and do not explain the various theories or competing perspectives in-depth. To give your essay greater analytical strength, you will need to go beyond the textbook and required readings, and demonstrate that you have read other relevant academic sources and are able to engage with theoretical perspectives and concepts in an intelligent way. Essays that provide evidence of in-depth/critical engagement with the academic literature (scholarly books, chapters in edited volumes and journal articles) are more likely to result in sound, critical and engaging analysis.

Different types of sources: **Primary** sources are documents created during the time in question by those who experienced the events. These may include newspapers, government reports, diaries, memoirs, United Nations reports, or reports by non-government organisations. Primary sources are written by the person who experienced the event, and therefore often contain biases, eg when a government justifies its policy position in the face of criticism. They tend not to be academic sources. They can be useful if your essay requires you to look at the attitudes of people directly involved in an event. **Secondary** sources are those in which people analyse the event in question. They can be scholarly (eg a scholarly journal article, a book or a textbook) or non-scholarly (eg a newspaper article).

Internet sources: You can use scholarly journals and texts which are available online. However, you should be wary of relying on other internet sources. Do not use google - use the University library resources or google scholar at <http://scholar.google.com>. Where you do use material taken from the internet, you should take particular care to check that it has been compiled by a dependable institutional source or by a reputable scholar. Remember that the purpose of researching an essay is to gather evidence pertinent to an argument and to demonstrate to the marker that you understand the different schools of thought or different contributions applicable to the topic.

Do not use these sources in writing an essay:

- online user-updated encyclopaedias such as *Wikipedia*
- blogs, newsgroups, email lists
- partisan, personal or anonymous websites.

It is always preferable to find the same information from a more reputable source. If you can, then use the reputable source as your reference. If you cannot, then it is best not to rely upon the information at all.

Newspapers and News Periodicals: These sources must be used cautiously, as they are not scholarly. They may provide you with facts, commentaries and occasional insight, but do not expect them to provide a coherent analytical framework. The Library has subscriptions to many national and international daily and weekly newspapers. Most are available online through *Factiva*.

Parliamentary and Government Records: Each Australian parliament records debates in *Hansard*. The federal parliament *Hansard*, as well as text of bills and other parliamentary information, can be found at: www.aph.gov.au.

6. Taking Notes

As you start taking notes, work with reference to your essay outline. There are many ways of writing down information from your sources. You may use a different set of notes for each of the sections of your essay outline. The major advantage of this system is that when you write your essay, each section of the outline can be dealt with in turn, without having to leaf through pages of disorganised notes.

Always write down the exact reference, including page number, for the information you write down in your notes. Always use quotation marks if it is a direct quotation, or paraphrase immediately into your own words, so as to avoid unintentional plagiarism. Using the arguments, ideas or words of another author, without acknowledging these via a citation or reference, is plagiarism. Plagiarism is a major violation of University rules and expectations.

Once you have finished your reading and taking notes, it is often useful to take a step back and think again about what you want to argue in the essay. Sometimes it helps to get some distance from the reading and your notes by taking a walk or waiting until the next day to revise the structure and argument of your essay or argument.

7. Writing and Revising

The essay should be in a **coherent and logical prose** that is cogently (convincingly and effectively) argued, carefully documented, and well written. The structure of the essay typically has three parts: an introduction, the body of the essay, and a conclusion.

- The **introduction** should introduce the topic to be discussed, state clearly your argument, and outline the points you will cover in the essay to make that argument – in the same order in which they will appear in the essay.
- The **body of the** essay is where the bulk of the argument is made. The body consists of a series of major paragraphs that are introduced with topic sentences, and developed in a logical sequence. Each paragraph should consist of one main idea. Introduce the main idea with a topic sentence, then provide evidence through data, examples, and the work of key scholars to support your idea. Reference all ideas or quotes that are not your own.
- The **conclusion** should restate briefly the key argument. You should show how your analysis – and your own distinctive approach – has allowed you to draw conclusions about the topic. Most often

the conclusion should be written in your own 'voice' and you generally should not include quotations from other authors in this section. Do not introduce new ideas at this stage.

There are two ways to get started when beginning the writing process. Either you can begin by writing your introduction first, or you can start with the body of the essay. Writing the introduction first may help to clarify the central argument of the essay, but remember that, like an essay plan, the introduction will often need to be revised as the essay progresses. The introduction can be the most difficult and time-consuming part of the essay to write. Some writers prefer to commence with the body of the essay, moulding the points from their outline into a structured argument. Once you have established the structure of the essay, this can serve as a guide for the introduction. Whichever method you use, the body of the essay must be consistent with the introduction and conclusion.

A very common failing is to write only one draft. If you wish to get the best possible mark on a university essay, you should allow sufficient time to write **more than one draft**. The aim of a first draft is to get the ideas mapped out on paper. One way of thinking about the first draft is as an 'expansion' of the essay outline. The aim of writing second (and subsequent) drafts is to refine your argument and to achieve the best possible wording. Expression can be corrected when writing subsequent drafts. When you are writing subsequent drafts you should also be focused on the cohesiveness of the essay. One way to revise a draft of your essay is to read it as if you were the marker. What are the shortcomings in the argument, writing, sequence, and so on? Revise the draft to overcome these deficiencies. Try to achieve an elegant writing style that you would enjoy reading.

Don't forget to give your paper a title; something that gives a brief and interesting 'snapshot' of your topic and argument.

In accordance with policies in regards to Student Integrity and Misconduct **you should not lend your original work to others** for any reason unless directed to by your Course Coordinator (for instance, if one of the tasks is a peer review). Nor should you collude with other people, including but not limited to fellow students, when completing your assessment work unless directed by your Course Coordinator (for instance, if you are undertaking a group essay).

Writing Conventions

Avoiding Bias: Bias refers to prejudices, preconceptions or predispositions that distort your capacity to examine and assess material in a dispassionate manner. It may be found in any of the following practices:

- ignoring or suppressing contradictory data or alternative views;
- using only writers who agree with your own viewpoint; or
- presenting dogmatic views or opinions that are not supported by evidence or argument.

The best way to avoid bias is to draw upon a broad range of sources and evaluate the arguments and assertions contained within them critically. It is also important to acknowledge the existence of alternative arguments and evidence to demonstrate the depth of your understanding to the marker. Readers will usually be expecting an essay to be explicit about different intellectual approaches to any particular theme.

Elegance of writing: Markers are always disappointed to read essays that display a considerable amount of research but are presented in an inelegant style. Lack of clarity in exposition is often a symptom of confused thinking. Here are some suggestions to improve your writing.

- **Role Models:** Pay attention to the style used in the articles and chapters that you read in researching your essay. Model your own writing style on the work of authors who you enjoy reading.
- **Grammar:** Inaccurate grammar and spelling distract the reader's attention from your ideas.
- **Dot points:** Do not submit an essay written in point form or with a series of one-sentence paragraphs. Write in complete sentences, with a verb.
- **Use active voice,** and avoid passive voice - it leads to long, complicated sentences. Compare the following: 'The bill giving the right to vote to women was passed by Parliament'; and 'Parliament passed the bill giving women the right to vote'. The second example uses active voice and is clear and straightforward.
- **Use the first person.** It is acceptable in the School for you to write in the first person. Rather than writing: 'In this essay it will be argued that...', try the alternative: 'In this essay I argue that ...'.
- **Quotations:** You should only use a quotation when the author's own words are critical to make the point you want to make. Otherwise, paraphrase the idea in your own words, remembering to give proper credit to the author. Quotations of **more than 40 words** should be presented as a 'block quote'; that is, they should be indented on both sides with single spacing in the text and presented without quotation marks. Always include the correct citation. Including the page number(s) for the quotation is essential.

Quotations must use the exact words and punctuation of the original text. If you want to omit some words from the middle of a quotation, indicate the omitted words with ellipses (...). If you want to add words or clarify comments, you must include them in square brackets []. For example, the quotation 'Never in the field of human conflict has so much been owed by so many to so few' could be shortened or clarified respectively as:

"Never in the field of human conflict has so much been owed ... to so few."

"Never in the field of human conflict has so much been owed by so many [people] to so few."

- **Gender-neutral language:** Avoid inappropriate gender-specific language, including gender-specific terms for groups of people or the characterisation of groups as male or female. The use of 'he', 'him', or 'his' as the default pronoun should not occur; do not use 'man' to mean humanity in general. Nor should you use female pronouns when referring to inanimate objects, for instance referring to a boat as she.
- **Non-racist language:** Terms that are discriminatory or prejudicial to ethnic or racial groups are unacceptable in academic writing. When referring to Australia's Indigenous peoples, the terms 'Aborigines and Torres Strait Islanders' should be used. 'Non-English speaking background' is used to denote someone whose cultural background is derived from a non-English-speaking tradition or whose first language is not English. Please use the full 'Lesbian, Gay, Bisexual, Transgendered, Intersex and Queer Community' when referring to this community. It is also suitable to use the acronym - LGBTIQ.
- **Cliches and jargon:** Avoid words and phrases that suffer from overuse. Clichés impede clear perception, feeling and thought. Phrases such as: 'the moment of truth', 'history tells us' and 'at this point in time' should be avoided.

- **Contractions, colloquialisms and slang:** Avoid contractions, slang and colloquial expressions. A contraction is a shorted form of a word or expression common in spoken English but should be avoided in formal written expression like essays. For instance use it is rather than it's and is not rather than isn't. A 'colloquialism' is a word or expression appropriate to a conversational level of usage, but is not suited to academic composition. Slang is a form of colloquialism where ordinary words have been given a special meaning; for example, words such as 'cool'.
- **Numbers:** Spell out the numbers one to nine and spell out even hundreds, thousands and millions. Use Arabic numerals (1, 2, 3, etc.) for other numbers. Percentages are expressed as figures followed by 'per cent' even if the number is less than 10 ('%' should be used only in charts, tables, graphs and footnotes). Always write out a number or year if it begins a sentence. Do not use an apostrophe if referring to a decade – for example, 1990s (**not** 1990's).
- **Acronyms:** An acronym is a word formed from the first (or first few) letters of a series of words. For example, AJPS is an acronym for the *Australian Journal of Political Science*. Acronyms should be in parentheses at the first reference, following the spelled-out full form. In later references the letters are sufficient:

Word count

Students are expected to write to the word limit set by the Course Coordinator. All words used in the text of your essay (including title, quotations, block quotations, in-text citations, tables, figures, headings) count as part of the word limit. Words used in the reference list, whether listed as a separate reference list at the end or in footnotes if you are using the footnoting system, do not count toward the word limit. Uploading your document as a 'Word' document on Turnitin will ensure the reference list is not counted in the calculation of total word count.

Essays and research reports are given a +/-10% leeway on the word count (unless the word count already stipulates a range). If an essay or research report exceeds this limit, it will attract a 10 percentage point penalty.

Presentation

Essays should be typed, use double or 1 ½ line spacing, have a margin of 2.5cm on all four sides of the page, use Times New Roman 12 point, or a similarly clear font and size, and should number all pages.

SECTION 2: REFERENCING

Which referencing system should you use?

The School **requires** that you use one of the two referencing systems set out in this Guide – either the author-date in-text referencing system as outlined in this Guide, or the footnote referencing system as outlined in this Guide. Use **one** of these two styles consistently and accurately.

Style 1: Author-date in-text referencing system

The citation in an author-date system consists of the last name of an author and the year of publication of the work, followed by the page numbers. Under this system, terms such as *ibid.* and *op. cit.* are **not** used. As a general rule, **citations must include individual page numbers.**

You will find variants of this referencing system in scholarly publications, including the Harvard system and the Chicago author-date system (http://www.chicagomanualofstyle.org/tools_citationguide.html).

To avoid confusion, please use the author-date in-text referencing system **as outlined in this Guide.**

Simple in-text citations with an identified author:

Australia's political culture has been characterised as consisting of compliant subjects rather than active citizens who genuinely participate in the country's civic life (Smith 2001: 27).

Jacobs initially advanced this idea (2002: 6), and it was later developed in the United States (Brown 2005: 92).

Ruling classes 'do not justify their power solely by de facto possession of it, but try to find a moral and legal basis for it' (Mosca 1939: 70).

More complicated instances of in-text citation:

- When **more than one study is cited**, in which case you separate the items with semi-colons, eg (Abato 2005: 34; North 2004: 256-260).
- When there are **two works by the same author**, in which case you separate each year by a comma, eg (Habermas 2000: 180-185, 2001).
- **When the same author has two or more sources with the same year**, in which case you give the first one a lower case 'a', the second one a lower case 'b', etc, eg (Vromen 2003a: 156-160, 2003b: 29-32).
- When there are **two or three named authors**, you put in both or all three names in the in-text citation, eg (Miragliotta, Errington and Barry 2013: 145).

- When there are **four or more named authors**, you write the first author's name and then put 'et al' after it, meaning 'and others', eg (Crenshaw et al 1995:345).
- When an **author's name is already listed in the sentence** that includes the in-text citation, you do not need to repeat the author's name in the in-text citation, eg Friedrich insists that 'constitutionalism, both in England and abroad, was at the outset not at all democratic' (1937: 31).
- **When an author cites another author**, you need to acknowledge this in the citation, eg A central question of politics has been formulated as 'who gets what, when and how?' (Lasswell 1936, cited in Dunleavy and O'Leary 1987: 18).
- Where **an author is not named**, you need to work out the institutional author, eg (Australian Bureau of Statistics 2014: 3), or (*The Economist* 2006: 12).
- Where you are **citing a reprint of an old book**, you need to put both the original publication date and the reprint date in the in-text citation, eg (Austen [1813] 2003: 57).
- When you are **citing a chapter in an edited book**, make sure you cite the author of *that* chapter, and not the editors of the book, in the in-text citation.
- When you are **citing legislation**, you follow the *Title Year* (Jurisdiction) format, eg The University is governed by the *University of Queensland Act 1998* (Qld), which requires ...
- When you are citing **UN documents**, you name the UN Department and year the document was agreed to in the in-text citation, eg (UNSC 2011: 3).

When using the author-date in-text citation system, footnotes are not used for citations, but can be used sparingly to expand on points in the text. Notes should be numbered consecutively and placed at the bottom of the page as footnotes. The corresponding note number in the text should be typed as a superscript.

Reference List/Bibliography

A reference list (also called a bibliography) of all the sources cited in the essay must be included at the end of your essay. The reference list should provide an accurate, alphabetically ordered by surname, and complete account of the sources you have cited in the essay. The reference list should not separate categories, eg books listed separately from journal articles. We have separated them here only to teach you how to list different sources. Where there are multiple sources by the same author in the same year, use the same lower case letter after the date you used in your in-text citation to differentiate between the sources.

Book

Gorard, Stephen 2003 *Quantitative Methods in Social Science* London: Continuum.

Multi-authored book

Singleton, Gwynneth; Don Aitkin; Brian Jinks and John Warhurst 2006 *Australian Political Institutions*, 8th ed. Melbourne: Longman.

Edited Book

Galligan, Brian ed. 1989 *Australian Federalism* Melbourne: Longman Cheshire.

Ivan Hare and James Weinstein eds. 2009 *Extreme Speech and Democracy* Oxford: Oxford University Press.

Chapter in an Edited Book

Beeson, Mark and Ann Capling 2002 'Australia in the World Economy', in S Bell ed. *Economic Governance and Institutional Dynamics* Melbourne: Oxford University Press.

Crenshaw, Kimberlé 1995 'Mapping the Margins: Intersectionality, Identity Politics, and Violence', in K Crenshaw, N Gotanda, G Peller and K Thomas eds. *Critical Race Theory: The Key Writings That Formed the Movement* New York: The New Press.

Scholarly Journal Article

Citrin, Jack; Eric Schickler and John Sides 2003 'What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections', *American Journal of Political Science* 47(1): 75-91.

Mackenzie, Chris 2004 'Policy Entrepreneurship in Australia: A Conceptual Review and Application', *Australian Journal of Political Science* 39(2): 367-386.

Newspaper and magazine where the author is named, and the source is hard copy

Stevenson, Richard 2006 'In Address, Bush Is Seen Avoiding Large Initiatives', *The New York Times* 26 January: 1.

Newspaper and magazine where the author is not named and the source is hard copy

The Australian 2006 'Europe Softens Stance on Iran', 19 January: 8.

Newspaper and magazine where the author is named, and the source is online

McGeough, Paul 2016 'Donald Trump to be an Unguided Missile at Debate. Will He Go Nuclear?' *Sydney Morning Herald* 10 October. Accessed 10 October 2016. Available at <http://www.smh.com.au/world/us-election/donald-trump-to-be-an-unguided-missile-at-debate-but-will-he-go-nuclear-20161009-gryf8y.html>.

Newspaper and magazine where the author is not named, and the source is online

Reuters 2016 'Myanmar Insurgents Kill at least 17 People in Targeted Attacks on Border in Rakhine State', *ABC News* 10 October. Accessed 10 October 2016. Available at <http://www.abc.net.au/news/2016-10-10/police-killed-in-myanmar-attacks-near-bangladesh-border/7917382>.

Online source that is only available online, not in hard copy

Bush, George 2005 *President's Address to the Nation*, 18 December. Accessed 25 January 2006. Available at <http://www.whitehouse.gov/news/releases/2005/12/20051218-2.html>.

Human Rights and Equal Opportunity Commission (HREOC) 2005 *Aboriginal and Torres Strait Islander Justice Report* Sydney: HREOC. Accessed 20 December 2005. Available at http://www.hreoc.gov.au/social_justice/index.html.

Government Report with named author

Manning, Ben and Roberta Ryan 2004 *Youth and Citizenship: A Report for the National Youth Affairs Research Scheme* Canberra: Department of Family and Community Services.

Government Report with institutional author

Department of Foreign Affairs and Trade (DFAT) 2005 *Annual Report 2004-2005* Canberra: AGPS.

United Nations Document

United Nations Security Council 2011 *Resolution 1970 (2011) Adopted by the Security Council at its 6491st Meeting on 26 February 2011, S/Res/1970*. United Nations, New York.

Translated Work

Politkovskya, Anna 2004 *Putin's Russia* (trans. Arch Tait) London: Harvil Press.

Reprints of older Work

Popper, Karl [1945] 2002 *The Open Society and Its Enemy*. London: Routledge.

Book Review

Weber, Jennifer 2006 'Andrew Johnson's Good Deed: Review of *Forever Free: The Story of Emancipation and Reconstruction*, by Eric Foner' *The Washington Monthly* 38(1): 50-52.

Kirchner, Stephen 2005 'Review of *Australia's Money Mandarins: The Reserve Bank and the Politics of Money*, by Stephen Bell', *Australian Journal of Political Science* 40(4): 567-568.

Unpublished Work

Hutchinson, Emma 2008 'Trauma, Emotion, and Political Community' PhD thesis, unpublished. St Lucia: University of Queensland.

Style 2: Footnote referencing system

There are multiple footnote referencing systems in use in scholarly publications. To avoid confusion, please use the footnote referencing system **as outlined in this Guide**, which is drawn from the Chicago notes referencing system.

The reference in a footnote referencing system uses superscript numbers, and footnotes are numbered consecutively through the essay. Each time a reference is needed, a superscript number appears, which links to a footnote at the bottom of the page in which the reference is placed. It looks like this.¹ The number usually appears at the end of the sentence, after any quotation marks, and after the full stop. Do not put a space in-between the full-stop and the number. The footnote itself lists the author's name in the order of first name and then surname.

When a footnote repeats a reference that has already appeared in a previous footnote, you need to create a sensible, brief version of the reference that appeared previously to identify which one you are referring to, but you do not need to repeat all of the reference material. As a general rule, **citations must include individual page numbers**.

When using this footnote referencing system, you must **also** provide a bibliography/reference list at the end of the essay. This is because the actual footnotes cite the actual page or page range that you are citing, and in the reference list you put the whole page range of the journal article or the book chapter. Also, in the footnotes if a book has four or more authors you use 'et al', and in the reference list you write the names of all the authors out in full.

Examples of footnotes

Books with one, two or three authors

1. Christian Reus-Smit, *The Moral Purpose of the State: Culture, Social Identity, and Institutional Rationality in International Relations* (Princeton, N.J.: Princeton University Press, 1999), 4-5.
2. Andrew Lynch and George Williams, *What Price Security: Taking Stock of Australia's Terror Laws* (Sydney: University of New South Wales Press, 2006), 145-150.

Books with four or more authors

3. Gwynneth Singleton et al, *Australian Political Institutions*, 8th ed. (Melbourne: Longman, 2006), 35-56.

Edited book

4. Brian Galligan ed. *Australian Federalism* (Melbourne: Longman Cheshire, 1989), 45-50.

Chapter in an Edited Book

5. Mark Beeson and Ann Capling, "Australia in the World Economy," in *Governance and Institutional Dynamics*, ed. Stephen Bell (Melbourne: Oxford University Press, 2002), 77.
6. Kimberlé Crenshaw, "Mapping the Margins: Intersectionality, Identity Politics, and Violence," in *Critical Race Theory: The Key Writings That Formed the Movement*, ed. Kimberlé Crenshaw, Neil Gotanda, Gary Peller and Kendall Thomas (New York: The New Press, 1995), 362.

Scholarly journal article

7. Jack Citrin, Eric Schickler and John Sides, "What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections," *American Journal of Political Science* 47(1) (2009): 75-91.

8. Chris Mackenzie, "Policy Entrepreneurship in Australia: A Conceptual Review and Application," *Australian Journal of Political Science* 39(2) (2004): 367-386.

Scholarly journal article that is available online with a doi

9. Vivien Lowndes, "Varieties of New Institutionalism: A Critical Appraisal," *Public Administration* 74(2) (1996): 410-432, accessed February 28, 2010, doi:10.1111/j.1467-9299.1996.tb00865.x.

Newspaper and magazine article where the author is named, and the source is hard copy

10. . Richard Stevenson, "In Address, Bush Is Seen Avoiding Large Initiatives," *New York Times*, January 26, 2006, 1.

Newspaper and magazine article where the author is not named and the source is hard copy

11. *The Australian*, "Europe Softens Stance on Iran," January 19, 2006, 8.

Newspaper and magazine article that you accessed online

12. Sheryl Gay Stolberg and Robert Pear, "Wary Centrists Posing Challenge in Health Care Vote," *New York Times*, February 27, 2010, accessed February 28, 2010, <http://www.nytimes.com/2010/02/28/us/politics/28health.html>.

Material sourced online that is only available online

13. George W. Bush, "President's Address to the Nation," December 18, 2005. Accessed 25 January, 2006. <http://www.whitehouse.gov/news/releases/2005/12/20051218-2.html>.

Government Report with named author

14. Ben Manning and Roberta Ryan, *Youth and Citizenship: A Report for the National Youth Affairs Research Scheme* (Canberra: Department of Family and Community Services, 2004), 99-101.

Government Report with institutional author

15. Department of Foreign Affairs and Trade, *Annual Report 2004-2005* (Canberra: AGPS, 2005), 99.

United Nations Document

16. . United Nations Security Council, *Resolution 1970 (2011) Adopted by the Security Council at its 6491st Meeting on 26 February 2011, S/Res/1970* (New York: United Nations, 2011).

Translated book

17. Gabriel García Márquez, *Love in the Time of Cholera*, trans. Edith Grossman (London: Cape, 1988), 242–55.

Book review

18. Jennifer Weber, “Andrew Johnson’s Good Deed,” review of *Forever Free: The Story of Emancipation and Reconstruction*, by Eric Foner, *The Washington Monthly* 38(1) (2006), 50-52.

19. Stephen Kirchner, review of *Australia’s Money Mandarins: The Reserve Bank and the Politics of Money*, by Stephen Bell, *Australian Journal of Political Science* 40(4) (2005), 567-568.

Unpublished Work

20. Emma Hutchinson, “Trauma, Emotion, and Political Community,” PhD thesis, unpublished. (St Lucia: University of Queensland, 2008).

Re-citing items that have already appeared in an earlier footnote

- 21. Reus-Smit, *The Moral Purpose of the State*, 26.
- 22. Lynch and Williams, *What Price Security*, 154.
- 23. Singleton et al., *Australian Political Institutions*, 60.
- 24. Galligan ed., *Australian Federalism*, 63.
- 25. Beeson and Capling, “Australia in the World Economy,” 79.
- 26. Crenshaw, “Mapping the Margins,” 375.
- 27. Citrin, Schickler and Sides, “What if Everyone Voted?” 94.
- 28. Mackenzie, “Policy Entrepreneurship,” 390.
- 29. Lowndes, “Varieties,” 414.
- 30. Stevenson, “In Address,” 2.
- 31. *The Australian*, “Europe Softens Stance,” 8.
- 32. Stolberg and Pear, “Wary Centrists.”
- 33. Bush, “President’s Address.”
- 34. Manning and Ryan, *Youth and Citizenship*, 103.
- 35. Department of Foreign Affairs and Trade, *Annual Report 2004-2005*, 101.
- 36. United Nations Security Council, *Resolution 1970*.
- 37. Márquez, *Love in the Time of Cholera*, 265.
- 38. Weber, “Andrew Johnson’s Good Deed,” 53.
- 39. Kirchner, review of *Australia’s Money Mandarin*, 569.
- 40. Hutchinson, “Trauma, Emotion,” 145.

Reference List/Bibliography

When using the footnote referencing style, you still need to provide a reference list at the end of the essay. This is because the items as they appear in the reference list contain the full source information, eg the page range of book chapters and journal articles and not only the page/s that you cited in the footnote.

The formatting for the reference list is different from the formatting you used in the footnotes. So please pay particular attention to the formatting in the reference list. For example, you include all your items in the reference list in alphabetical order with the surname listed first, then the first name (which is the opposite of the order you used in the footnotes themselves), and you do not use brackets around the publishers, place of publication and year.

Please use the following format in the reference list:

- Beeson, Mark and Ann Capling. "Australia in the World Economy." In *Governance and Institutional Dynamics*, edited by Stephen Bell, 77-90. Melbourne: Oxford University Press, 2002.
- Bush, George W. "President's Address to the Nation." December 18, 2005. Accessed 25 January, 2006. <http://www.whitehouse.gov/news/releases/2005/12/20051218-2.html>.
- Citrin, Jack, Eric Schickler and John Sides. "What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections." *American Journal of Political Science* 47(1) (2009): 75-91.
- Crenshaw, Kimberlé. "Mapping the Margins: Intersectionality, Identity Politics, and Violence." In *Critical Race Theory: The Key Writings That Formed the Movement*, edited by Kimberlé Crenshaw, Neil Gotanda, Garry Peller and Kendall Thomas, 357-383. New York: The New Press, 1995.
- Department of Foreign Affairs and Trade. *Annual Report 2004-2005*. Canberra: AGPS, 2005.
- Galligan, Brian. ed. *Australian Federalism*. Melbourne: Longman Cheshire, 1989.
- Hutchinson, Emma. "Trauma, Emotion, and Political Community." PhD thesis, unpublished. St Lucia: University of Queensland, 2008.
- Kirchner, Stephen. Review of *Australia's Money Mandarins: The Reserve Bank and the Politics of Money*, by Stephen Bell. *Australian Journal of Political Science* 40(4) (2005): 566-568.
- Lowndes, Vivien. "Varieties of New Institutionalism: A Critical Appraisal." *Public Administration* 74(2) (1996): 410-432. Accessed February 28, 2010. doi:10.1111/j.1467-9299.1996.tb00865.x.
- Lynch, Andrew and George Williams. *What Price Security: Taking Stock of Australia's Terror Laws*. Sydney: University of New South Wales Press, 2006.
- Mackenzie, Chris. "Policy Entrepreneurship in Australia: A Conceptual Review and Application." *Australian Journal of Political Science* 39(2) (2004): 362-386.
- Manning, Ben and Roberta Ryan. *Youth and Citizenship: A Report for the National Youth Affairs Research Scheme*. Canberra: Department of Family and Community Services, 2004.
- Márquez, Gabriel García. *Love in the Time of Cholera*. Translated by Edith Grossman. London: Cape, 1988.
- Reus-Smit, Christian. *The Moral Purpose of the State: Culture, Social Identity, and Institutional Rationality in International Relations*. Princeton, N. J.: Princeton University Press, 1999.
- Singleton, Gwynneth, Don Aitkin, Brian Jinks and John Warhurst. *Australian Political Institutions*, 8th ed. Melbourne: Longman, 2006.
- Stevenson, Richard. "In Address, Bush Is Seen Avoiding Large Initiatives." *New York Times*, January 26, 2006.
- Stolberg, Sheryl Gay and Robert Pear. "Wary Centrists Posing Challenge in Health Care Vote." *New York Times*, February 27, 2010. Accessed February 28, 2010. <http://www.nytimes.com/2010/02/28/us/politics/28health.html>.
- The Australian*. "Europe Softens Stance on Iran." January 19, 2006.
- United Nations Security Council. *Resolution 1970 (2011) Adopted by the Security Council at its 6491st Meeting on 26 February 2011, S/Res/1970*. New York: United Nations, 2011.
- Weber, Jennifer. "Andrew Johnson's Good Deed." Review of *Forever Free: The Story of Emancipation and Reconstruction*, by Eric Foner. *The Washington Monthly* 38(1) (2006): 50-55.

INTRoDucTIoN to

SOCIAL RESEARcH

QUANTITATIVE & QUALITATIVE APPROACHES

KEITH F PUNCH

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC

List of tables	xiii
List of figures	xv
Preface	xvii
Companion website	xix
About the author	xxi
1 Introduction	1
2 Theory and Method in Social Science Research	13
3 Ethics in Social Science Research	35
4 Research Questions	57
5 From Research Questions to Data	73
6 Literature Searching and Reviewing	93
7 Qualitative Research Design	113
8 Collecting Qualitative Data	143
9 Analysing Qualitative Data	167
10 Quantitative Research Design	205
11 Collecting Quantitative Data	227
12 Analysing Quantitative Data	251
13 The Internet and Research	279
14 Mixed Methods and Evaluation	301
15 Research Writing	327
Glossary	347
Appendix 1	355
References	359
Index	377

fifteen

RESEARCH WRITING

Contents

- 15.1 Background
- 15.1 The quantitative tradition
- 15.1 Qualitative research writing
- 15.1 The analytical mix
- 15.2 Research documents
- 15.2 Proposals
- 15.2 Qualitative proposals
- 15.2 Mixed methods proposals
- 15.2 Examples of proposals
- 15.2 Abstracts and titles
- 15.2 Dissertations (and projects)
- 15.3 Writing to report versus writing to learn: writing as analysis
- 15.4 Writing choices
- Chapter summary
- Exercises and study questions
- Further reading

Learning objectives

After studying this chapter you should be able to:

- Describe the typical structure of a quantitative research report
- Explain how this structure has broadened for qualitative research writing
- List headings showing the content expected in a proposal, and briefly describe what each heading indicates
- List headings showing the content expected in a dissertation, and briefly describe what each heading indicates
- Explain what is meant by the internal consistency of a proposal or dissertation

Writing is an important part of research. Getting a project started usually means taking it from ideas to a written proposal. At the other end, a project is not complete until it is shared, through writing. Thus, a written proposal is required for the project to commence, and a written report is required after the project. It follows that the quality of research is judged in part by the quality of the written document (proposal or report).

The first part of this chapter gives some background to the topic of research writing, by looking briefly at writing in the quantitative tradition and at the much greater range of writing choices in qualitative research, and by using the 'analytical mix' as a device for thinking about quantitative and qualitative writing together. The second part then deals in some detail with research proposals and, in less detail, with abstracts and dissertations. Final sections then discuss briefly the distinction between writing to report and writing to learn and some of the choices facing the social science researcher when it comes to writing.


Background

5.1.1 The quantitative

The conventional format for reporting (or proposing) quantitative research has such headings as these (Miles, Huberman and Saldana, 2013):

- statement of the problem;
- conceptual framework;
- research questions;
- method;
- data analysis;
- conclusions;
- discussion.

A still briefer form used by some journals (for example, journals with a strong behaviourist leaning in psychology and education) has just four headings – introduction



and research questions, method, results, discussion. For much quantitative research reporting, this framework of headings is still quite appropriate – Gilbert (2008), for example, describes the conventional (quantitative) sociological paper with similar headings. But many qualitative researchers today would find such headings too constraining. The headings are still relevant to much of what we do, but we often have different and broader expectations, for a qualitative research report especially, than can be met by these headings. Once again this reflects the range of perspectives in qualitative research in contrast with the relative homogeneity of much quantitative research.

2.1.2 Qualitative

Throughout this book I have used the timing-of-structure continuum (Section 2.6) to stress the range of social science research, especially qualitative research. Writing about quantitative research has typically been a relatively straightforward matter, with conventional models and structures such as those in the previous section to guide the writer. Writing for qualitative research, however, like the research itself, is much more varied and diverse, and not at all monolithic (Coffey and Atkinson, 1996). Thus, towards the right-hand end of this continuum, there is a greater range of writing models, strategies and possibilities. Some of the perspectives of contemporary qualitative research (such as feminism and postmodernism, especially in ethnography) broaden the range even further.

The rethinking of research which has accompanied both the paradigm debates and the emergence of new perspectives has included research writing – how the research is to be put into written form and communicated. Especially when seen in a discourse analysis or sociology of knowledge perspective, this rethinking has brought awareness of the choices about writing, identifying the conventional quantitative writing model as just one of the choices. The appreciation of a wider range of choices has meant also the freeing up of some of the restrictions about writing, and the encouraging of experimentation with newer forms of writing.

As a result, there is a proliferation of forms of writing in qualitative research, and older models of reporting are being mixed with other approaches:

The reporting of qualitative data may be one of the most fertile fields going; there are no fixed formats, and the ways data are being analyzed and interpreted are getting more and more various. As qualitative data analysts, we have few shared canons of how our studies should be reported. Should we have normative agreement on this? Probably not now – or, some would say, ever. (Miles, Huberman and Saldana, 2013)

2.1.3 The

In this chapter I discuss research writing for both quantitative and qualitative approaches, keeping in mind but not overemphasising their differences, and keeping in mind also mixed methods studies that combine both approaches. Adapting Miles

TABLE 15.1 Terms in the analytic mix

Quantitative	Qualitative
Variable-oriented	Case-oriented
Categorising	Contextualising
Analytical	Synthetic
Etic	Emic
Variance theory	Process theory

and Huberman's 'analytic mix' (1994: 301-2) is a helpful device in doing this, since it brings key elements of the two approaches together.

One part of this analytic mix was used in Chapter 14 (pp. 292-5), where variable-oriented research (quantitative) was compared with case-oriented research (qualitative). As Table 15.1 shows, other writers have used slightly different terms.

Miles and Huberman write about this mix from within qualitative research, expressing strongly the view that good qualitative research and reporting requires both types of elements from the mix. We can use it as a framework for all research writing.

The framework echoes the two tensions within qualitative research described by Nelson et al. (1992) in discussing cultural studies, and paraphrased by Denzin and Lincoln (2011). On the one hand, qualitative research is drawn to a broad, interpretive, postmodern, feminist and critical sensibility. On the other, it is drawn to more narrowly defined positivist, post-positivist, humanistic and naturalistic conceptions of human experience and its analysis. These latter conceptions correspond more to the left-hand side of Table 15.1, whereas the former conceptions are closer to the right-hand side.

Miles and Huberman note that Vitz (1990) reflects these two views, seeing conventional data analysis, involving propositional thinking, as the fruit of abstract reasoning, leading to formal, theoretical interpretations. Figural genres, such as narratives, entail more concrete, holistic reasoning; they are 'stories' retaining the temporal configurations of the original events.

When it comes to writing, Miles and Huberman recommend a mixture of these two ways of seeing the world. As they see it, the problem is that:

Stories without variables do not tell us enough about the meaning and larger import of what we are seeing. Variables without stories are ultimately abstract and unconvincing - which may explain certain scrupulous rules for reporting quantitative studies, as well as the familiar comment, 'I couldn't really understand the numbers until looked at the open-ended data'. (1994: 302)

The challenge therefore is to:

combine theoretical elegance and credibility appropriately with the many ways social events can be described; to find intersections between the propositional thinking of mass

conventional studies and more figurative thinking. just as good analysis nearly always involves a blend of variable-oriented, categorizing, 'paradigmatic' moves, and case-oriented, contextualizing, narrative ones, so does good reporting. (1994: 299)

These ideas of the range of styles, and of mixing and blending them, are a useful background against which to consider proposals, abstracts and dissertations.

Research documents

There are two main types of research documents - proposals and reports (which may be dissertations, journal articles or reports for other purposes). Since this book is an introduction to social science research, with much focus on getting the research started, most attention here is on proposals. Much of what is said carries over naturally to dissertations and articles. Abstracts and titles are also briefly considered in this section.

2.1 Proposal

This section summarises my views on research proposals. A full description of these views is given in Punch (2006), where all of the points made below are described in detail, and where five exemplary proposals are presented in full.

What is a research proposal? In one sense, the answer is obvious - the proposal describes what the proposed research is about, what it is trying to achieve and how it will go about doing this, and what we will learn from this and why it is worth learning. In another sense, the dividing line between the proposal and the research itself is not so obvious. The proposal describes what will be done, and the research itself is carried out after approval of the proposal. But preparing the proposal itself may also involve considerable research.

The three most basic questions are useful in guiding development of the proposal:

What? What is the purpose of this research? What are we trying to find out?

How? How will the proposed research answer these questions?

Why? Why is this research worth doing (or funding)? Or, what will we learn, and why is it worth knowing?

The first question (what?) is dealt with in Chapters 4 and 5. The second question (how?) is dealt with in Chapters 7, 8 and 9 for qualitative research, and Chapters 10, 11 and 12 for quantitative research. The third question (why?) is discussed later in this section.

Maxwell (2012) stresses that the form and structure of the proposal are tied to its purpose - 'to explain and justify your proposed study to an audience of non-experts on your topic'. *Explain* means that your readers can clearly understand what you want to do. *Justify* means that they not only understand what you plan to do,

but why. *Your proposal should be mainly about* study, not mainly about the literature, your research topic or research methods in general. *Nonexperts* means that researchers will often have readers reviewing their proposals who are not experts in the specific area.

It is helpful to see the proposal itself as an argument. Seeing it as an argument means stressing its line of reasoning, its internal consistency and the interrelatedness of its different parts. It means making sure that the different parts fit together, and showing how the research will be a piece of disciplined inquiry, as described in Chapter 14. As an argument, it should explain the logic behind the proposed study, rather than simply describing the study. In so doing, it should answer the question of why this design and method is chosen for this study.

What follows now is a suggested set of guidelines for developing a research proposal, shown in Table 15.2. Because no prescription is appropriate, but certain content is expected, and because there are both similarities and differences in quantitative and qualitative proposals, it seems best to present a full checklist of possible sections and headings. Some clearly apply to both quantitative and qualitative research, whereas some are more applicable to one approach. Not all would necessarily be required in any one proposal, and they can be used as appropriate. They are things to think about in proposal preparation and presentation, and they are useful in developing full versions of the proposal – where shorter versions are required, a good strategy is to prepare the full version, then summarise it.

Many of these headings (for example, general and specific research questions) do not now require comment, because of what has already been said in previous chapters. Where this is the case, the reader is referred to appropriate parts of the book. Where new points need to be made, or earlier ones reinforced, or where important distinctions apply between quantitative and qualitative approaches, brief comments are made. I stress that these are suggested guidelines. As with writing a report or dissertation, there is no fixed formula for a proposal. There are different ways to present the material, and different orders the sections can follow.

It is easier in many respects to suggest proposal guidelines for a quantitative study, since there is greater variety in qualitative studies, and many qualitative studies will be unfolding rather than prestructured. An emerging study cannot be as specific in the proposal about its research questions, nor about details of the design. When this is the case, the point needs to be made in the proposal. A discussion of qualitative proposals follows this section (see Section 15.2.2).

There are many ways a topic can be introduced, and all topics have a background and a context. These need to be dealt with in the introduction, which sets the stage for the research. A strong introduction is important to a convincing proposal. Its purpose is not to review the literature, but rather to show generally how the proposed study fits into what is already known, and to locate it in relation to present knowledge and practice. In the process of doing this, there should be a clear identification of the research area and topic, and a general statement of purpose. This can

Checklist of possible sections for research proposals

Title and title page
Abstract
Introduction: area and topic
 background and context
 statement of purpose
Research questions: general and specific
Conceptual framework, theory, hypotheses (if appropriate)
The literature
Methods: strategy and design
 sample and sampling
 data collection - instruments and procedures
 data analysis
Significance
Consent, access and human participants' protection
References
Appendices (e.g. timetable, budget, instruments, etc.)

then lead into the research questions in the next section. Specific features of the proposed study can also be identified here, as appropriate – for example, if personal knowledge or experience form an important part of the context, or if preliminary or pilot studies have been done, or if the study will involve secondary analysis of existing data (Maxwell, 1996).

For qualitative proposals, two other points may apply here. One is the first general evaluative question given in Chapter 14 – What is the position behind this research? If this question is applicable, it can be answered in general terms, to orient the reader early in the proposal. The other is more specific – Where on the structure continuum is the proposed study? This strongly influences later sections of the proposal. If a tightly structured qualitative study is planned, the proposal can proceed along similar lines to the quantitative proposal. If a more emergent study is planned, where focus and structure will develop as the study proceeds, this point should be made clearly. In the former case, there will be general and specific research questions. In the latter case, there will be only general orienting research questions.

Research questions

These were discussed in detail in Chapters 4 and 5. In the proposal outline suggested, they can follow from the general statement of purpose given in the introduction.

Theory and

There is wide variation in the applicability of this section. If it applies, it is a matter of judgement whether the conceptual framework goes here, or in the methods

section later in the proposal. Theory and hypotheses are included if appropriate as explained in Chapter 4. If theory is involved, it may be included in the literature review section, rather than here. The role of theory in the research should be made clear here, however.

Literature

The proposal needs to be clear on the position taken with respect to the literature in the proposed study. As discussed in earlier chapters, there are three main possibilities:

The literature is reviewed comprehensively in advance of the study, and this review is included as part of the proposal.

The literature will be reviewed comprehensively ahead of the empirical stage of the research, but this review will not be done until the proposal is approved. In this case, the nature and scope of the literature to be reviewed, and familiarity with it, should be indicated.

The literature will deliberately not be reviewed prior to the empirical work, but will be integrated into the research during the study, as in grounded theory research. In this case too, the nature and scope of the literature should be indicated.

For some qualitative proposals, the literature may be used in sharpening the focus of the study, and to give structure to its questions and design. If so, this should be indicated, along with how it is to be done. In all cases, the researcher needs to connect the proposed study to the literature (see, for example, Marshall and Rossman, 1989; Locke et al., 1993; Maxwell, 1996).

Strategy and design In all proposals, whether quantitative, qualitative or mixed methods, it is a good idea to start this section with a short clear paragraph describing the strategy the research will use for answering the research questions (see the example given at the end of Section 14.4.4, p. 313). The basic quantitative designs we have discussed are the experimental, quasi-experimental and correlational survey designs. For these designs or variations of them, the conceptual framework may be shown here, instead of earlier. In qualitative studies, the location of the study along the structure continuum is particularly important for its strategy and design. As noted in Chapter 7, qualitative designs such as case studies (single or multiple, cross-sectional or longitudinal), ethnography or grounded theory may overlap, and elements of these strategies may be used separately or together. This means it will be difficult to compartmentalise the study neatly. This is not a problem, but it should be made clear that the proposed study uses elements of different strategies. Qualitative studies vary greatly on the issue of predeveloped conceptual frameworks, and the position of the proposed study on this matter should be indicated. A fully or partly predeveloped framework should be shown.

Where one will be developed, it needs to be indicated how this will be done. This will internet with data collection and analysis, and may be better dealt with there. Mixed methods studies should identify, describe and justify which mixed methods design is proposed.

Sample As shown in Chapter 11, the three key sampling issues for quantitative research are the size of the sample, how it is to be selected and why, and what claims are made for its representativeness. The qualitative proposal should deal with the questions of who or what will be studied, and why. The sampling strategy is important for both types of studies, and especially important for both the quantitative and qualitative parts of a mixed methods study, and its logic needs to be clear. Where the sampling strategy itself is emergent, as in theoretical sampling, this needs to be explained.

Data collection The two matters here are the instruments (if any) that will be used for data collection, and the procedures for administering the instruments. If a quantitative study proposes to use instruments that already exist, and information about their psychometric characteristics is available, it should be included. If the instruments are to be developed, the general steps for developing them should be shown. If a qualitative study proposes to use instruments (for example, observation schedules, structured or semi-structured interviews), the same comments apply. Less structured qualitative data collection techniques should be indicated and discussed, especially in terms of the quality of data issues shown in Section 14.5.3. For quantitative, qualitative and mixed methods studies, the procedures proposed for data collection should also be described, and the description should show why these data collection activities have been chosen. Possible threats to the validity of data, and strategies to minimise or control these threats, can also be indicated here.

Data analysis Quantitative proposals should indicate the statistical procedures by which the data will be analysed. Similarly, the qualitative proposal needs to show how its data will be analysed, and how the proposed analysis fits with the other components of the study (see Section 9.8, p. 200). A mixed methods proposal needs to cover both types of analysis. If applicable, all types of proposal should indicate what computer use is planned in the analysis of the data.

The particular topic and its context will determine the study's significance. Other terms for this might be 'justification', 'importance' or 'contribution' of the study – they all address the third earlier overarching question: Why is this study worth doing? There are three general areas for the significance and contribution of the study – to knowledge in the area, to policy considerations and to practitioners (Marshall and Rossman, 1989). The first of these, contribution to knowledge, is closely tied to the literature in the area. One function of the literature review is to indicate gaps in the knowledge in the area, and to show how this study will contribute to filling those gaps. This has to be set against the position taken on the literature, as discussed above.

These are dealt with in Chapter 3 on ethical issues.

This is a list of the references cited in the proposal.

These may include any of the following: a timetable for the research, letters of introduction or permission, consent forms, measuring instruments, questionnaires, interview guides, observation schedules, examples of pilot study or other relevant work already completed (Maxwell, 2012).

15.2.3.0.0

Qualitative studies vary greatly, and in many, the design and procedures will evolve. This obviously means that the writer cannot specify exactly what will be done in advance, in contrast with many quantitative proposals. When this is the case, there is a need to explain the flexibility the study requires and why, and how decisions will be made as the study unfolds. Together with this, as much detail as possible should be provided. Review committees have to judge both the quality, feasibility and viability of the proposed project, and the ability of the researcher to carry it out. The proposal itself, through its organisation, coherence and integration, attention to detail and conceptual clarity, can inspire confidence in the researcher's ability to execute the research. In addition, where specialised expertise is involved (for example, advanced statistical analysis, or grounded theory analysis), it helps for the researcher to indicate how this expertise will be acquired.

Marshall and Rossman (1989) stress the need for the qualitative proposal to reassure the reader as to the academic merit and discipline of the proposed research. This need is less pronounced today, when there is very much greater recognition and acceptance of qualitative research. However, there are two main ways the qualitative proposal can provide this reassurance. One is by giving information about the technical issues of the research, under research methods, as is routinely done in quantitative proposals; this means the sampling plan, the data collection and quality of data issues, and the proposed methods of analysis. The other applies to an unfolding qualitative study. Its proposal should indicate that focus will be developed as the study proceeds, and how this focus will be developed during the early empirical work.

Contrasting design and proposals at different ends of the structure continuum in qualitative research, Denzin and Lincoln (1994: 200) write:

The positivist, postpositivist, constructionist, and critical paradigms dictate, with varying degrees of freedom, the design of a qualitative research investigation. This can be broken

at as a continuum, with rigorous design principles on one end and emergent, less well-structured directives on the other. Positivist research designs place a premium on the early identification and development of a research question and a set of hypotheses, choice of a research site, and establishment of sampling strategies, as well as a specification of the research strategies and methods of analysis that will be employed. A research proposal may be written that lays out the stages and phases of the study. These phases may be conceptualised ... (reflection, planning entry, data collection, withdrawal from the field, analysis, and write-up). This proposal may also include a budget, a review of the relevant literature, a statement concerning protection of human subjects, a copy of consent forms, interview schedules, and a timeline. Positivist designs attempt to anticipate all of the problems that may arise in a qualitative study. Such designs provide rather well-defined road maps for the researcher. The scholar working in this tradition hopes to produce a work that finds its place in the literature on the topic being studied.

In contrast, much greater ambiguity is associated with postpositivist and nonpositivist designs – those based, for example, on the constructivist or critical theory paradigms or the ethnic, feminist, or cultural studies perspectives. In studies shaped by these paradigms and perspectives there is less emphasis on formal grant proposals, well-formulated hypotheses, tightly defined sampling frames, structured interview schedules, and predetermined research strategies and methods and forms of analysis. The researcher follows a path of discovery, using as a model qualitative works that have achieved the status of classics in the field.

Thus, for some types of qualitative research especially, we do not want to constrain too much the structure of the proposal, and we need to preserve flexibility. On the other hand, several writers (Coffey and Atkinson, 1996; Silverman, 2011) point out that this should not be taken to mean that 'anything goes'. Eisner (1991: 241-2) writes in the same vein about qualitative research in education:

Qualitative research proposals should have a full description of the topic to be investigated, a presentation and analysis of the research relevant to that topic, and a discussion of the issues within the topic or the shortfalls within the research literature that make the researcher's topic a significant one. They should describe the kinds of information that are able to be secured and the variety of methods or techniques that will be employed to secure such information. The proposals should identify the kinds of theoretical or explanatory resources that might be used in interpreting what has been described, and describe the kinds of places, people, and materials that are likely to be addressed.

The function of proposals is not to provide a watertight blueprint or formula the researcher is to follow, but to develop a cogent case that makes it plain to a knowledgeable reader that the writer has the necessary background to do the study and has thought clearly about the resources that are likely to be used in doing the study, and that the topic, problem, or issue being addressed is educationally significant.

Lest these comments be interpreted by some to mean that no planning is necessary in conducting qualitative research, or that 'anything goes', as they say, I want to make it

clear that this is not how my words should be interpreted. Planning is necessary. Nevertheless, it should not and cannot function as a recipe or as a script. Evidence matters. One has a responsibility to support what one says, but support does not require measured evidence. Coherence, plausibility, and utility are quite acceptable in trying to deal with social complexity. My point is not to advocate anarchy or to reduce the study of schools and classrooms to a Rorschach projection, it is to urge that the analysis of a research proposal or a research study should employ criteria appropriate to the genre. Professors who make such assessments should understand, as should graduate students, the nature of the genre, what constitutes appropriate criteria, and why they are appropriate.

15.2.3 Mixed methods proposals

Once both the research questions and the design to answer them are clear in a mixed methods study, a strategy-design statement such as that shown in Section 14.4.4 can be constructed for the proposal. After this (and depending on the design chosen), it is often convenient to split a mixed methods proposal into its quantitative and qualitative parts, and to describe the sampling, data collection and data analysis for each part. However, while this splitting can help in presentation, it is important also to show the connections between the two parts. This is especially true for sampling – for example, when a qualitative in-depth second stage follows a larger sample quantitative first stage. After describing the sampling strategy for the first stage, an important question is: How will the second stage subsample be chosen and why?

1 Examples of proposals

The literature contains some useful examples of research proposals. In addition to the five shown in Punch (2006), a detailed treatment of proposals of different types is given by Locke et al. (1993). They present four examples of research proposals in full, and they give a detailed critical commentary on the different sections and aspects of each. They have chosen the proposals to illustrate different designs and styles of research, and using topics drawn from different areas. The four proposals are:

An experimental study ('The effects of age, modality, complexity of response and practice on reaction time'). This study proposes a two-factor design with repeated measures, to test 14 hypotheses about reaction times.

A qualitative study ('Returning women students in the community college'). This study proposes the use of in-depth interviewing to explore the meaning of experiences of older women returning as students to a community college.

A quasi-experimental study ('Teaching children to question what they read: An attempt to improve reading comprehension through training in a cognitive learning strategy'). This study proposes a quasi-experimental design to test three hypotheses about the acquisition of reading by children.

A funded grant proposal ('A competition strategy for worksite smoking cessation'). This renewal grant proposal also uses a quasi-experimental design, to assess the effectiveness of competition/facilitation on recruiting employees into a self-help smoking cessation programme, and on the outcomes of that programme.

In addition to these examples, Maxwell (2012) presents a qualitative proposal entitled 'How basic science teachers help medical students learn: The students' perspective', and he too gives a detailed commentary on the proposal. The research he describes proposes to use a case study of four exceptional teachers to answer six specific research questions about how teachers help medical students learn. Classroom participant observation and student and teacher interviews are the main sources of data, supplemented by relevant documentary data. Finally, Chenitz (1986) does not include an example of an actual proposal, but writes about the preparation of a proposal for a grounded theory study.

An abstract is a brief summary, whether of a proposal or a finished study. Abstracts play an important role in the research literature, and they are required in proposals (usually), in dissertations and in research articles in refereed journals. Abstracts and titles are at the heart of the hierarchical indexing system for the research literature, which becomes more and more important as the volume of research continues to build. This indexing system enables researchers first to scan a title, to see if they need to go further into a project. If so, they can go to the abstract, which will tell them more, and perhaps enough. If they still need to go further, the last chapter (for example, of a dissertation) will often contain a summary of the study and its findings, in more detail than the abstract. They can then go to the full report if they need still more detail about the research.

Good abstract writing requires the skill of saying as much as possible in as few words as possible. For a proposal, the abstract needs to deal with two main issues -what the study is about and aims to achieve (usually best stated in terms of its research questions), and how it intends to do this. For a report, the abstract would need three main sections -these two, and a third which summarises what was found. The abstract should give an overview not just of the study itself, but also of the argument behind the study, and this should run through these sections. For most of us, abstract writing is a skill that needs to be developed, since we typically use many superfluous words when we speak and write. Together with the title, the abstract is usually written last, since it is difficult to summarise what has not yet been written.

Titles also have importance in the research literature indexing process, as indicated. Therefore a title should not just be an afterthought, nor should it use words or phrases that obscure rather than reveal meaning. Extending the point about abstract writing, the title should convey as much information as possible in as few words as possible. Titles and their role are discussed by Locke et al. (1993).

As noted, much of the focus in this book has been on getting research started, so the emphasis in this chapter is on writing the research proposal. Completed research is reported in several forms, and dissertations are one of the main forms. Because I have emphasised proposals, there is not a detailed description of the dissertation here, nor guidelines for its structure and writing. There is a considerable literature on this topic, and directions into this literature are given in the suggestions for further reading at the end of the chapter. Instead, this section now includes comments about three aspects of a dissertation – about the general content a dissertation should cover, about how a dissertation might be seen, and about the nature of dissertation writing.

Whatever its specific chapter structure, certain basic content is expected in a dissertation, which forms the report of a piece of research. This content includes:

- clear identification of the research area and the topic;
- a statement of purpose(s) and research questions;
- a setting of the study in context, including its relationship to relevant literature;
- a description of methods, including strategy and design, sample and the collection and analysis of data;
- a presentation of the data and of its analysis;
- a clear statement of the findings and a consideration of what can be concluded from those findings.

These headings are general enough to cover quantitative and most qualitative work. They are similar to Miles and Huberman's minimum guidelines for the structure of a qualitative research report (1994: 304):

The report should tell us what the study was about or came to be about.

It should communicate a clear sense of the social and historical context of the setting(s) where data were collected.

It should provide us with what Erickson (1986) calls the 'natural history of the inquiry', so we see clearly what was done, by whom and how. More deeply than in a sheer 'methods' account, we should see how key concepts emerged over time; which variables appeared and disappeared; which codes led into important insights.

A good report should provide basic data, preferably in focused form (vignettes, organised narrative, photographs, or data displays) so that the reader can, in parallel with the researcher, draw warranted conclusions. (Conclusions without data are a sort of oxymoron.)

Finally researchers should articulate their conclusions, and describe their broader meaning in the worlds of ideas and action they affect.

How the material is divided up into chapters and sections is a matter of judgement for the dissertation writer. In making this judgement, it is useful to remember that a dissertation is essentially the report of a piece of research, and the research itself constitutes a logical argument. Empirical research (quantitative, qualitative or mixed methods) systematically introduces empirical evidence into this argument, as

its way of answering questions, testing hypotheses or building understanding. In line with this, one way to look at research, including dissertation research, is as a series of decisions. Especially in planning and designing the project, the researcher faces choices, many of which have been the subject of this book. Therefore the completed project itself is a combination of these choices, and the dissertation is the report of this. It is very often not the case that there is a right and wrong choice in the face of these many decisions. As I have stressed frequently in this book, it is rather a case of assessing each situation in the research, along with its alternative choices and their inevitable strengths and weaknesses, and of making each decision based on this analysis, in the light of the circumstances of the research, and of the need for the parts of the project to fit together.

To reflect this perspective in the writing of the dissertation, the writer can say what the choices were at each point, what choice was made and why. Seeing a dissertation this way makes it clear that there is no one way to do a piece of research, and that any piece of research will have its critics. Recognising this, the objective is to produce a thorough report of a carefully reasoned set of consistent choices, after consideration of the alternatives. In the written report, the writer is, among other things, telling the reader about the decision path taken through the research, and taking the reader down this path. The writing indicates why this path was chosen, the alternatives considered and the decisions taken. Presenting it this way conveys the impression of a thorough and careful project, well planned, well executed and well reported.

Much has been written on the topics of tactics and style in dissertation and academic writing, and both topics are covered in the further reading indicated. The following comments concern the need for clarity, the role of shortening, and the modular and iterative nature of research writing.

Whether proposal or report, research writing needs to communicate clearly and effectively, and striving for clarity is part of the writer's responsibility. Clarity is required in the structure of the document (the sections it will have, the order in which they appear, and how they are connected to each other), and in the words, sentences and paragraphs that make up the sections. Clear guidelines help in these matters in quantitative research, but it is rather more difficult to balance clarity with 'fidelity' in the qualitative context. Quoting Berger and Kellner (1981), Webb and Glesne point out that we have a moral obligation to reflect stories of human meanings as faithfully as possible. Reflecting the style of qualitative analysis emphasised in Chapter 9 of this book, with abstractions from the data to first order and then to second order constructs, they write (1992: 804):

The second order constructs that social scientists use to make sense of peoples' lives must grow from and refer back to the first order constructs those same people use to define themselves and fill their lives with meaning. Moving gracefully between first and second order constructs in writing is difficult. We have suggested that students have a moral obligation to be clear; otherwise the reader is separated from the lives of the people under study and from the researcher's analysis. We are not suggesting that students should dumb-down their text or simplify what is complex just to make reading easier.

One general strategy to help in being clear is to put oneself (as writer) in the position of reader. What will make most sense to the reader? What will ensure that the reader is easily able to follow the argument and does not 'get lost' in the document? In this way, the writer tries to anticipate reader expectations and reactions. It is useful to remember also that the research document (proposal or report) will, in the end, be a stand-alone document. This means that the writer will not be there to interpret it for the reader, when it is being read.

Shortness is important because it often promotes clarity. There are various pressures today to restrict length – thus research journals restrict article length because of space considerations, and universities place upper limits on the length of dissertations. Ways to achieve shortness include getting straight to the point, cutting out unnecessary padding, not using long words when short ones will do, and keeping sentences short. The problem is that shortening requires reworking, which takes time. Every researcher discovers that lack of time is a problem in making the writing short (and clear). The message is to leave adequate time to do the shortening, if possible.

The organisation and structure of a dissertation (or proposal) usually require that it be segmented into sections. It is useful to see these as modules, to organise these modules into chapters, and to write the dissertation by writing these modules. Breaking it up makes the task less formidable. It is also helpful to write the different sections, or at least to keep full notes and draft the sections, as the stages of the research are being carried out. So many issues and decisions arise during a research project that it is impossible to remember all of them when it comes to 'writing time' without full notes of the various discussions, readings and so on. Wolcott (1990) goes even further with his advice: 'You cannot begin writing early enough.' The strategy of 'writing as you go' helps in making ideas explicit (Miles, Huberman and Saldana, 2013), and it also exploits the value of writing as a part of the learning. Some specific strategies to help in academic writing are given in Punch (2006: 72-4).

In the traditional model of research writing, the write-up does not get done until the research is completed and everything is figured out. 'I've done all the research, now I am writing it up.' Implicit in this is the idea that I don't start the writing until I have 'got it all worked out'. This is *writing to report*.

A different view sees writing as a way of learning, a way of knowing, a form of analysis and inquiry. This is the idea of 'writing in order to work it out'. In this view, I don't delay the writing until I have it all figured out. On the contrary, I use the process of writing itself to help me figure it out, since I learn by writing. Writers interpret, so writing is a way of learning, through discovery and analysis (Richardson, 1994). Thus writing becomes an integral part of the research, and not just an add-on once the 'real' research is completed. This is *writing to learn*.

Writing to learn is more likely in qualitative research. However, it can also have a role in quantitative studies – for example, when the researcher is interpreting the results from the analysis of complex data sets, as in a multivariable correlational survey. In these cases, building an overall picture of the results, integrating and interpreting them, is similar to describing the emerging picture which accompanies some forms of qualitative data analysis (such as the Miles and Huberman type). At the same time, this model of writing is especially appropriate for some types of qualitative analysis, where the researcher is constructing a map or theoretical picture of the data, which emerges as the analysis proceeds. The process of writing can be a great help in developing this emerging picture. A practical implication of this view is that a useful tactic, when stuck or confused in developing the picture, is to attempt some writing about it.

Qualitative researchers are therefore more likely to stress that writing is analysis, not separated from it. In the 'analytic work of writing' (Coffey and Atkinson, 1996), writing is part of thinking, analysing and interpreting. The 'crisis of representation' (Denzin and Lincoln, 1994) has also brought shifts in conceptions of how to represent the 'reality' with which qualitative research (especially) has to deal, particularly the world of lived real experience. Together, these two points bring a new focus on the form of the written research report:

The net effect of recent developments is that we cannot approach the task of 'writing up' our research as a straightforward (if demanding) task. We have to approach it as an analytical task, in which the form of our reports and representations is as powerful and significant as their content. (Coffey and Atkinson, 1996: 109)

This view, particularly prominent in recent writing about ethnography (Hammersley, 1995; Coffey and Atkinson, 1996), leads to a realisation of the many choices involved in the production of research documents.

Webb and Glesne (1992: 803) order the writing choices facing the researcher, especially the qualitative researcher, on three levels. Macro questions are about power, voice and politics in research; middle-range issues concern authorial authority, the marshalling of evidence, and the relationship between the researcher and the researched; micro issues are about whether a piece should be written in the first person, whether its tone changes when the author moves from data to theory, and how the story is told. Miles, Huberman and Saldana (2013) also identify a series of choices about reports and reporting, stressing choices rather than a fixed set of ideas. They include choices about the report's audiences and its hoped-for effects on them, the voice or genre of the report, its writing style, and its structure and format. To go with this array of choices, Denzin and Lincoln add that the process of writing is itself an interpretive, personal and political act (2011).

These writing choices in fact apply across the whole range of research approaches quantitative, qualitative and mixed methods, though it is the developments qualitative research in the past 30 years which have demonstrated this most clearly. The further reading indicated below, especially in qualitative research writing, includes discussions and examples of these choices.

Summary

The quantitative tradition gives us a straightforward set of headings for writing research proposals and reports; by contrast, qualitative research has introduced a much broader range of writing models and structures.

Research proposals need to address the overarching questions of what, how and why, and a set of headings is presented to show the proposal as an argument, with a good fit between its component parts, while answering these questions.

Qualitative proposals may require greater flexibility, and in some respects are more difficult to write, especially if an unfolding study is proposed.

In writing mixed methods proposals, it is often useful to separate the methods section into qualitative and quantitative parts, but an overall description of strategy and design should show how the parts interconnect to answer the research questions.

Abstracts and titles are important in indexing the research literature, and require the skill of conveying as much information as possible, in as few words as possible.

Seeing the dissertation as the report of a piece of research leads to a clear view of the content it is expected to contain.

'Writing to report' is typical of the quantitative tradition in research; 'writing to learn' is often used in a qualitative context, where writing is seen as part of the analysis and inquiry.

What is meant by the quantitative tradition in research reporting? What is the typical structure of a report, in this tradition?

How and why has qualitative research broadened writing strategies for proposals and dissertations?

What three central questions guide proposal development? Discuss each question, and how each can be dealt with in a proposal.

What does it mean to say that a research proposal is an argument?

What sections would you expect to find in a research proposal, and what is the function of each?

What chapters would you expect to find in a dissertation, and what is the function of each?

What is a research abstract, and why is it important?

Study a recent edition of a research journal such as the *American Educational Research Journal*, *The Administrative Science Quarterly*, the *British Journal of Psychology* or *The American Sociological Review*. What do you learn about research writing by studying the titles, abstracts and structure of the articles?

Broadley, L. (1987) *Academic Writing in Social Practice*. Philadelphia: Temple University Press.

Mullins, C.I. (1977) *A Guide to Writing and Publishing in the Social and Behavioural Sciences*. New York: Wiley.

Strunk, W. and White, E.B. (1979) *The Elements of Style*. 3rd edn. New York: Macmillan.

Zinsser, W. (1976) *On Writing Well*. New York: Harper and Row.

Atkinson, P. (1990) *The Ethnographic Imagination: Textual Constructions of Reality*. London: Routledge.

Becker, H.S. (1986) *Writing for Social Scientists: How to Finish Your Thesis, Book, or Article*. Chicago: University of Chicago Press.

Clifford, J. and Marcus, G.E. (1986) *Writing Culture: The Poetics and Politics of Ethnography*. Berkeley, CA: University of California Press.

Geertz, C. (1983) *Works and Lives: The Anthropologist as Author*. Cambridge: Polity Press.

Van Maanen, J. (1988) *Tales of the Field: On Writing Ethnography*. Chicago: University of Chicago Press.

Wolcott, H.F. (1990) *Writing Up Qualitative Research*. Newbury Park, CA: SAGE.

Glatthorn, A. and Joyner, R. (2005) *Writing the Winning Dissertation: A Step-by-Step Guide*. London: SAGE.

Krathwohl, D.R. (1988) *How to Prepare a Research Proposal*. 3rd edn. Syracuse, NY: Syracuse University Press.

Locke, L.F., Spirduso, W.W. and Silverman, S.J. (1993) *Proposals That Work: A Guide for Planning Dissertations and Grant Proposals*. 3rd edn. Newbury Park, CA: SAGE.

Long, T.J., Convey, J.J. and Chwalek, A.R. (1991) *Completing Dissertations in the Behavioural Sciences and Education*. San Francisco: Jossey-Bass.

Madsen, O. (1992) *Successful Dissertations and Theses*. 2nd edn. San Francisco: Jossey-Bass.

Phillips, M. and Pugh, D.S. (1987) *How to Get a PhD*. Milton Keynes: Open University Press.

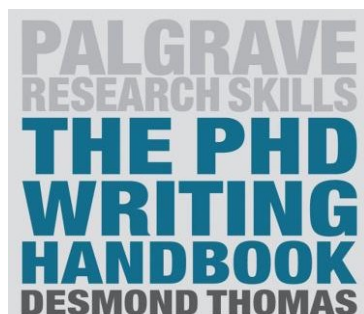
Punch, K.F. (2006) *Developing Effective Research Proposals*. 2nd edn. London: SAGE.

Rudestam, K.E. and Newton, R.R. (2000) *Surviving your Dissertation: A Comprehensive Guide to Content and Process*. 2nd edn. Thousand Oaks, CA: SAGE.

Sternberg, O. (1981) *How to Complete and Survive a Doctoral Dissertation*. New York: St Martin's.

Thody, A. (2006) *Writing and Presenting Research*. London: SAGE.

Walliman, N. (2004) *Your Undergraduate Dissertation: The Essential Guide to Success*. London: SAGE.



The PhD Writing Handbook

1st edition

Desmond Thomas

Macmillan Education UK

Paperback 9781137497697 £20.99 / \$42.00

Do you feel overwhelmed by the prospect of writing your thesis? Are you concerned that your PhD will take over your life? This much-needed text will help you to develop the writing skills you need to succeed at every stage of your research. It covers core topics such as:

- developing a research topic through writing
- exploring key concepts through writing
- building a structured chapter framework
- completing a first draft
- writing clearly, coherently and in an appropriate academic style

Each chapter features insights from researchers along with tasks and self-evaluation exercises to help you develop your own strategies for success. Invaluable to thesis writers from a range of disciplines, *The PhD Writing Handbook* will help you to think more clearly and communicate more effectively with your readers.

Contents

<i>List of Illustrative Materials</i>	viii
<i>Acknowledgements</i>	xi
Introduction	1
1 The Stages of PhD Thesis Writing	4
The traditional model of the PhD	4
Alternative models of the PhD	7
Making a start	11
2 Developing a Research Topic through Writing	15
Choosing a topic	15
Finding and developing a topic that is feasible	17
Finding and developing a topic that is specific and significant	22
Developing an updated rationale for the research	31
3 Managing Research Reading	34
Reading aims	34
Problem areas associated with reading	37
Notetaking: From reading into writing	47
4 Producing a Literature Review	55
The literature review as a process and a product	55
The purposes of a literature review	56
Alternative approaches to the literature review as product	58
Literature review problem areas	63
5 Exploring Key Concepts	67
The importance of concept development within a research project	67
The starting point for concept exploration	69
The process of concept exploration: the next steps	72
Other ways of exploring concepts through lateral thinking	75
Exploring concepts in writing	77

vi Contents

6	Building a Structured Chapter Framework	80
	Producing an overall framework	80
	Types of chapters within the structural framework	88
	Chapter introductions and conclusions	91
7	Establishing Productive Writing Routines	100
	Goals, schedules and routines	100
	When will you write? How will you allocate time?	102
	Setting goals and targets	107
	Where and how will you write?	107
	Feedback on writing	110
8	Writing Clearly, Concisely and Coherently	116
	Clarity and coherence	116
	The nature of coherence	122
	Coherence at chapter level	126
9	Developing Academic Style	128
	Expectations of academic style	128
	Expectations of difficulty	131
	Classifying 'difficult texts'	134
10	Completing a First Draft	141
	The basic writing cycle	141
	Managing reader expectations	142
	Drafting and redrafting	145
	Monitoring progress	146
	Writing and editing	150
11	Motivating Yourself to Write	155
	Writing is more than just a technical exercise	155
	The nature of writer's block	156
	In which situations does writer's block occur?	158
	What are the causes of writer's block?	162
	What kind of strategies or solutions can we use for writer's block?	165
	Coping with rejection	168
12	Reporting and Analysing Research Findings	171
	Structuring your data report	171
	Reporting quantitative data	175
	Reporting qualitative data	177
	Drawing conclusions	181

13	Completing your PhD Thesis	185
	Writing the final version	185
	Oral presentations and writing	193
	The viva	201
14	Writing throughout an Academic Career	205
	<i>Appendix 1: Useful Websites</i>	211
	<i>Appendix 2: Suggested Answers for Task 2.1</i>	213
	<i>Appendix 3: Research Rationale Template</i>	215
	<i>References</i>	217
	<i>Index</i>	222

Introduction

A PhD research project presents two types of challenges. The first relates to the project itself, the type of research that is being carried out, its content and its methodology. For these, there are many potential sources of support – individual supervisors, supervisory panels, discipline-related communities and an ever-expanding range of courses and books that offer training and advice on research design, data collection and data analysis.

The other set of challenges relates to the ability to produce a written PhD thesis that will do justice to the quality of your research, enabling readers to understand and appreciate the development of your thinking. This ability will also sustain you throughout a future career in academia, when writing will be the main means of communication with other members of disciplinary and cross-disciplinary communities. It is evident that developing writing skills to the required level will be an important goal for all PhD researchers. However, there is a tendency for writing support to be made available only through occasional self-contained workshops while the need for ongoing support is often underestimated or even ignored.

There are several possible reasons for the relative lack of attention to writing which can characterize research training programmes in higher education institutions. One is the fact that many PhD supervisors and research skills trainers see their responsibilities as limited to giving guidance on content and methods. This is a great pity, as most supervisors and trainers tend to be experienced and skilled writers within their specialist fields. It seems that expertise in writing is rarely passed on from the experienced researcher to the less experienced in an explicit way.

Another reason for the lack of writing support may be an underlying belief that writing is a simple and inherently logical process, which involves working out and clarifying ideas before you write them down. From this perspective, there is a designated order of events where writing comes after reading, collecting data and analysing data as the final stage of PhD research. From this same perspective, writing does not need to be taught. Instead, it is believed that the skills involved can be acquired over a period of time. I have known many cases where PhD researchers were forbidden

2 Introduction

to write by their supervisors until their third or fourth year of study on the grounds that they might not be ready.

For inexperienced research writers, acquiring the range of skills that they need can present a substantial challenge. The realization that writing a first draft can be a complex and messy process may come very gradually. The stereotype of the inspired writer spontaneously producing a brilliantly clear and coherent text can persist despite strong evidence to suggest that this view of writing rarely matches the reality.

One way of challenging the stereotype is to consider the role of writing and its importance at every stage of the research process. According to Wolcott (1990), writing is not really a product of thinking but is arguably a form of thinking in itself. Putting ideas down on paper or on a screen can be an excellent way of clarifying key issues, helping the writer to see arguments and counterarguments more clearly or examine the strength of the evidence presented to support a particular point that is being made. From this perspective, some form of writing needs to take place even during the very early stages of research, where writing contributes to thinking and helping to make reading more focused.

This handbook seeks to convince you that such an approach to writing can provide a useful basis for developing the skills that you will need to communicate with your intended readers. Some form of writing is present at every stage of the PhD research cycle: this has always been the case and perhaps always will be.

The handbook is divided into sections which follow the process of PhD thesis writing from the initial research proposal up until the final viva.

Chapter 1 considers different models of the PhD and issues related to reading and writing that need to be addressed at the start of PhD research, while *Chapter 2* considers ways of developing a suitable research topic through writing.

Chapters 3, 4 and 5 look at various aspects of reading in order to write: the management of the reading process, issues related to compiling a literature review and the exploration of key concepts.

Chapter 6 considers ways of building a structured chapter framework for the thesis, while *Chapter 7* focuses on the importance of establishing productive writing routines which also include receiving and responding to feedback from various sources.

Chapters 8, 9 and 10 break down the process of first draft writing, examining in turn issues related to clarity and coherence at sentence, paragraph and chapter level, academic style conventions and the drafting and redrafting of text.

Chapter 11 discusses some effective ways of achieving self-motivation and self-discipline in writing, having considered the causes and consequences of writer's block.

Chapters 12 and 13 consider the later stages of thesis writing, including the reporting and analysing of data. Practical guidelines are provided for the effective and systematic editing of written text, the development of oral presentation skills and preparation for the viva.

Finally, *Chapter 14* considers ways in which writing skills can be developed throughout an academic career.

Index

A

abstract

- final version of 187
- first draft of 84
- guidelines for writing an 85
- as a guide to text selection 41

abstract nouns 118, 130, 138

academic

- apprenticeship 4
- conventions (*see* disciplinary norms/conventions *and also* style conventions)
- publications 9–11, 205
- style (*see* style)
- writing (*see* writing)

allocating time for writing

(*see under* writing)

ambiguity of meaning 83, 118

analytical summaries 53–4

annotating a reading text 51–3

(*see also* notetaking)

appendices 65, 178, 193

assumed knowledge 64–5

B

bibliography 192

blogging 149–50

brainstorming 28

C

case studies 180

categorization of reading texts

- impossible texts 40

key/important texts 39–40

possible texts 40

seminal texts 39

single item texts 40

useful texts 40

chapters

concluding sections of 96–7

headings/subheadings, wording of 86, 88

introductions to 91–5

number of 86

sequence of/order of 83, 86–7

structure of 86–7

themes, identifying/

highlighting 95–6

chapter types

conclusion chapters 90–1, 188

introductory chapters 89

literature review chapters (*see under* literature review)

methodology chapters 83, 89–90

reporting/analysing findings 90, 171–183

claims/hypotheses

assumptions underlying 45

disagreeing with 45–6

formulating 28–9

hedging 30–1, 129–30, 191

qualifying through limiting

conditions 29–30

testing 62

clarity

at sentence level 117–20

principles 119
 problems of 117–21
 co-authorship 10–11, 207
 coherence
 across cultures 125–6
 at paragraph level 120–2, 143–5
 at chapter level 126–7, 146
 at whole text level 44–5, 152–3
 definitions 123–4
 of argumentation 44–5
 principles 120
 problems 121–2, 126–7
 the nature of 122–4
 vs structure 122–4
 commentary, analytic 10
 complex structures 117–8, 136, 138, 191
 concepts
 contested 63, 77
 dictionary definitions of 69
 developing/defining 67–79
 fluidity of 68–9
 guides for developing 78–9
 origins of 77, 79
 concept exploration
 based on experience/intuitive understanding 69–71
 considering associations/relationships 71, 75–6
 considering derivations 76
 degrees of separation 76
 different modes of 78
 finding antonyms 75
 in writing 77–8
 placing on a scale 75–6
 through diagrams 71, 76
 through lateral thinking (*see* lateral thinking)
 unpacking pairs/groups of 69, 72, 78
 core literature 36
 counterarguments 125, 152
 criteria for selecting reading texts (*see* categorization of reading texts)

D

daily writing schedules
 (*see under* writing)
 data
 access 18
 analysis categories 173
 breakdowns 176–7
 collection methods 172
 display formats 176–9
 drawing conclusions from 181–3
 highlighting key findings in 172, 175–6, 181–3
 layers 181
 reporting checklist 188
 reports, structuring of 172–5
 quantitative, reporting of 175–7
 qualitative, reporting of 177–81
 transcripts 167, 178, 193
 verbatim 178–9
 difficult texts
 containing complex subject matter 134
 debate concerning 131–4
 idiosyncratic style 134–5
 intentionally difficult 136–8
 mixed styles 139
 simplification of 138–40
 translated 135–6
 digressions 122
 disciplinary norms/conventions 59–60, 86–8, 120, 129–31, 137, 139, 171, 173–4
 discourse markers 123–4

E

editing (*see under* writing)
 editing checklists
 coherence, clarity and style 190–2
 content 186–189
 proofreading 192–3
 structure 189–90
 elegance (*see* stylishness)
 emotive language 46, 50

224 Index

- ethical
 - approval 16, 21
 - considerations/guidelines 20-1
- evaluating a research topic (*see under* feedback)
- exaggeration 46, 50
- expectations of readers
 - anticipating 142-3, 171
 - clarifying 13-14, 56, 58, 131-4, 151
 - managing 81, 142-5, 177
- F**
- feasibility of research topics (*see under* topic for research)
- feedback
 - from editors 207
 - from peers 111-2, 149-50, 207
 - from supervisors 112-4, 127, 151
 - negative, responding to 114, 148-9, 151, 169-70, 207
 - self-evaluation/self-monitoring 66, 110-11, 146-50, 209-10
- footnotes 65, 136, 190
- H**
- hedging (*see under* claims/hypotheses)
- hierarchy of ideas (*see under* PhD thesis)
- I**
- illustrative narratives 95, 139
- J**
- jargon 133-4, 137-8, 207
- journal, reflective writing 146-9, 166
- L**
- language accuracy/problems 152, 163-4
- lateral links in research 23-6
- lateral thinking
 - puzzles 73-5
 - techniques 72, 74-6
- linking expressions (*see* discourse markers)
- literature review
 - across disciplines/alternative approaches 58-60
 - as process 55-8
 - as product 55-6, 58-60
 - checklist 187
 - interim 60
 - methodology 59-60
 - problem areas 63-6
 - purposes 56-8
 - structure 62-3
 - thematic organization of 63
 - voice, in a 37, 61-2, 65
- M**
- mediation document 84, 100, 103
- mind maps 24-6, 109
- monitoring progress (*see under* feedback)
- motivation for writing 155-6, 159-60, 207
- N**
- narrative of thinking 41
- notetaking
 - level 1/surface 47-50
 - level 2/critical 50-3
 - sequence of 50-1
 - techniques for 47-53, 166
 - using audio recordings for 108-9
- O**
- obscurity 119, 136
- oral presentations
 - communication skills/delivery 196-7
 - content of 194-5

criteria for 194–9
 dealing with audience
 questions 199–200
 research in progress 199–200
 role of 193–4
 structure/coherence 196–8
 tips for 200–1

P

paragraphs/paragraphing 41–2, 120,
 143–5, 190
 paraphrasing 48–9
 passive, use of the 118–9, 130
 perfectionism 157, 163, 207
 PhD
 alternative models 7–11,
 by publication/by papers 9–11, 58,
 64, 87
 Creative 8–9
 New Route 7, 100
 professional doctorates 7–8, 28
 quitting a 161–2
 stages 4–7
 thesis framework (*see under* PhD
 thesis)
 traditional model of 4–7
 writing tasks 6
 PhD thesis
 chapters (*see* chapters)
 hierarchy of ideas in 86
 ownership of 160–1
 sections, examples of 91
 statements/statements of
 intent 61, 121–2
 structural framework 5, 80–91
 title 85–6
 polishing text (*see under* editing
 checklists proofreading)
 portfolio of work 7–8
 proposal, initial (*see under* rationale
 for research)

Q

qualifying research claims (*see under*
 claims/hypotheses)
 quotations
 excessive/inappropriate use
 of 134, 139
 in chapter introductions 93–4
 multiple/block 179
 single 178–9

R

rationale for research
 initial 11–13
 structure of 32
 templates 11–2, 215–6
 updated 31–2
 readership 63–4, 208
 reading
 access to sources 38
 across cultures 46–7
 aims 34–5
 aloud 122
 critica 144–6
 early stages of/exploratory 35–
 6
 establishing criteria for text
 selection (*see* categorization of
 reading texts)
 focused 36–7
 intensive 39–40, 110
 later stages of 36–7
 leaving footprints 40–1
 level 1/surface 43–4
 level 2/critical 44–6
 non-critical/surface 43–4
 notetaking for (*see under*
 notetaking)
 problem areas 13, 35, 37–47,
 110
 strategies 43–4
 redundancy 137–8, 191

226 Index

- referencing
 - backward 92
 - forward 97
 - in-text 128, 153, 192
- rejection, coping with 168–70, 207
- repetition 128, 137, 191
- research
 - context 32
 - distance 19–20, 166
 - findings (*see under* data)
 - hypotheses, formulation of (*see under* claims)
 - initial proposal 11–12
 - practical applications of 28
 - problem/underlying question 27–8
 - process, reflecting on 183, 209–10
 - questions, formulation of 25–7, 36
 - theoretical basis 18–19
 - topic (*see* topic for research)
- researcher
 - experiences 9, 11, 58, 107, 160
- resilience 156
- rhetoric 116, 133
- S
- scanning 44
- scare quotes 135
- self-awareness 16, 146–7, 165–8 (*see also* expectations)
- self-confidence 159–60, 163–4
- self-discipline 147
- signposting 65, 143–5, 190
- significance of research topics (*see under* topic for research)
- simplifying texts 118, 138
- skimming 41, 43
- specificity of research topics (*see under* topic for research)
- style
 - conventions 117, 128–30, 137
 - vs clarity 116, 120, 131–4, 136–8,
 - stylishness/elegance 130–1
 - summarizing 34–5, 48
 - supervisors (*see under* feedback)
 - synthesizing 49
 - T
 - tables of contents 81–4, 97–8
 - terminology (*see* concepts)
 - thematic narrative 174
 - thesis (*see* PhD thesis)
 - topic for research/research study
 - breaking down a 22–6
 - choosing a 15–17
 - developing a 4–5
 - evaluating a 21–2
 - feasibility 17–22
 - justifying a 57–8
 - originality 26–7
 - significance 26–8, 57
 - specificity 22–31
 - topic sentences 41–3, 125, 143–5
 - V
 - vague language 46, 118–9
 - vignettes 180–1
 - viva/viva voce examination
 - participating in 204
 - preparing for 202–3
 - role of 201–2
 - W
 - writer
 - responsibility 125
 - stance 43–4
 - stereotypes 2
 - voice 37, 122, 130 (*see also* literature review, voice in a)
 - writer's block
 - causes of 162–5
 - nature of 156–8
 - situations 158–62
 - solutions/strategies 165–8
 - writing
 - and editing 119, 150–3, 185–6
 - and editing checklists (*see under* editing checklists)

- as thinking 2, 148, 208
- cycle 141–2, 150
- engaging with 108–10, 206
- experience 206–8
- first draft 141–50, 152
- freely 150–2
- journal (*see* journal, reflective writing)
- locations 107–8
- monitoring of (*see under* feedback)
- plans (monthly) 102–4
- principles 5
- problem areas 5, 10, 13, 111, 148
- productive 101–2
- re-drafting of 141–2, 145–6,
- routines 100–10, 155
- schedules (daily) 104–6
- setting goals/targets for 107
- skills 206
- starting 159
- support 1–2, 106
- the final version 185–6
- throughout an academic career 168–9, 205, 208–9
- time management 102–6, 168

LITERATURBERICHT

Politische Kommunikation im Umbruch – neue Forschung zu Akteuren, Medieninhalten und Wirkungen

Gerhard Vowe / Marco Dohle

1. Einführung*

Kein anderer Tag hat die Welt der Gegenwart stärker geprägt als der *11. September 2001*. Im Licht dieses Ereignisses haben sich auch die Konturen politischer Kommunikation schärfer denn je abgezeichnet. Insbesondere ist deutlich geworden, wie sich Medien und politische Akteure wechselseitig beeinflussen: wie sie einander beobachten, benutzen und beschränken. So konnte die Welt durch Medienaugen nicht nur die Anschläge selbst *beobachten*, sondern auch wahrnehmen, wie die politischen Akteure darauf reagierten. Und umgekehrt konnten Politiker über die Medien beobachten, wie sich die öffentliche Meinung zu den Anschlägen überall in der Welt bildete und veränderte. Politische Akteure verschiedenster Art versuchten rund um diesen 11. September die Medien zu *benutzen*, um ihre jeweiligen Ziele zu erreichen. So zielten bereits die Anschläge darauf, globale Ikonen zu schaffen. Und ein gewichtiger Teil der Gegenmaßnahmen sollte diesen Bildern andere Bilder entgegensetzen – bis hin zu Bildern von der Demütigung des Feindes. Dies geschah in der Überzeugung, dass Medien politisch etwas bewirken – im Denken, Wollen und Tun der Menschen. Umgekehrt benutzten auch die Medien die politischen Akteure, wenn sie aus deren Kampf und Streit den Stoff destillierten, der die Menschen berührt und erregt. Und schließlich ist in der Folge des 11. Septembers deutlich geworden, wie die beiden Seiten einander *beschränken*: Die US-Regierung hat feste Regeln gesetzt, wie die Journalisten aus den Kriegen in Afghanistan und im Irak zu berichten haben. Andererseits geben die Medien den politischen Akteuren vor, was Nachrichtenwert hat und was nicht, wer glaubwürdig ist und wer nicht, wann der richtige Zeitpunkt für eine Inszenierung ist und welche Dramaturgie sie haben muss.

Der 11. September 2001 hat den Stellenwert von Kommunikation für die Politik sichtbar gemacht. Zugleich ist damit deutlich geworden, welcher Stellenwert einer Forschung zukommt, die das Konzept der Kommunikation zum Ausgangspunkt für den Zugang zur Politik wählt. Dieser Zweig der Forschung wird im Folgenden resümierend dargestellt; dabei soll nicht zuletzt erkennbar werden, welchen Beitrag deutschsprachige Publikationen leisten.

* Die Autoren danken R. Schmitt-Beck, Ch. Holtz-Bacha, O. Jarren, St. Marschall und W. Schulz für Kommentare zu einer früheren Fassung dieses Literaturberichts.

Für ein überschaubares Resümee sind in mehrfacher Hinsicht *Einschränkungen* erforderlich. Erstens wird das Schwergewicht darauf gelegt, die in die Zukunft weisenden Linien der Forschung herauszuarbeiten, also zu markieren, wo sich in der Literatur der letzten vier Jahren innovative Themen, Befunde und Methoden abzeichnen. Der gut abgesicherte Stand der Forschung ist in den mittlerweile zahlreichen Lehr- und Handbüchern zur politischen Kommunikation zu finden (aus politikwissenschaftlicher Sicht: Sarcinelli 2005; Strohmeier 2004; aus kommunikationswissenschaftlicher Sicht: Jarren/Donges 2006; Kaid 2004; Lilleker 2006; McNair 2003; Schulz 2007; fachübergreifend: Jarren et al. 2002; die ältere Literatur in der Übersicht: Schulz 1997; systematische Inhaltsanalysen von Forschungsliteratur: Graber 2005; Vowe/Dohle 2006). Dieser kognitive Grundbestand ist – neben spezialisierten Fachzeitschriften wie „Political Communication“ – ein klares Indiz dafür, dass der Forschungsbereich mittlerweile etabliert ist. Ebenso unberücksichtigt bleiben alle Formen unvermittelter politischer Kommunikation – von diplomatischen Verhandlungen bis hin zum Lobbying (Kleinfeld et al. 2007). Die Konzentration auf mediale politische Kommunikation kann aus einer Grundannahme der Forschung begründet werden: Politische Kommunikation habe sich „mediatisiert“ – die Orientierung an den Medien präge die Kommunikation in der Politik insgesamt (Pontzen 2006; Rössler/Krotz 2005; Vowe 2006). Des Weiteren muss außen vor bleiben, welcher politische Handlungsbedarf aufgrund von Veränderungen der politischen Kommunikation gesehen wird. Fragen der Medienpolitik oder der „Media Governance“ im nationalen und internationalen Maßstab (Donges 2007; Holtz-Bacha 2006) können hier nicht behandelt werden.

Der Bericht ist nach den drei Gesichtspunkten politischer Kommunikation gegliedert, bei denen die Veränderungen in der Forschung besonders klar hervortreten. Zunächst geht es darum, wie die Akteure der politischen Kommunikation gesehen werden; im nächsten Schritt, wie Veränderungen der Inhalte und Formen politischer Kommunikation gefasst werden; und schließlich, wie politische Kommunikation im Hinblick auf Nutzung und Wirkung untersucht wird. Unter allen drei Aspekten wird deutlich werden, wie die Forschung einen grundlegenden Wandel der politischen Kommunikation in den letzten Jahren nachzuvollziehen versucht:

- die Erweiterung der tradierten Akteurskonstellation politischer Kommunikation,
- die Abkehr von der Fixierung auf Information und Wort bei der Untersuchung der Medieninhalte sowie
- die Entdeckung und Nutzung neuer Wirkungspotentiale.

2. Veränderungen der politischen Kommunikation unter dem Akteursaspekt

Die *Akteurskonstellation* der politischen Kommunikation wird gewöhnlich als ein Dreieck aus politischen Organisationen, Medienorganisationen und der Bürgerschaft dargestellt, in dem sich zwischen den einzelnen Akteuren stabile Kommunikationsbeziehungen herausgebildet haben (z. B. McNair 2003: 6).¹ Als Ausgangspunkt für die Charakterisierung der Akteure wird zumeist die jeweilige Interessenlage genommen: Politische

¹ Ein Beispiel für einen nicht akteurstheoretisch basierten Ansatz in der Forschung zur politischen Kommunikation ist Japp/Kusche 2004.

Organisationen wie Parteien, Verbände, staatliche Instanzen, internationale Organisationen und NGOs konkurrieren um politischen Einfluss und nutzen die Möglichkeiten der politischen Kommunikation, um in dieser Konkurrenz Vorteile zu erlangen. Medienorganisationen wie Agenturen, Verlage, Fernsehsender oder Online-Anbieter konkurrieren um Anteile auf den verschiedenen Medienmärkten und nutzen ihrerseits die politische Kommunikation, um in dieser Konkurrenz Vorteile zu erhalten, die es ihnen erlauben, ihren Gewinn oder ihr Budget zu maximieren. In ihrer politischen Rolle als Bürger sind die Individuen daran interessiert, mit möglichst geringem Aufwand möglichst großen Einfluss auf politische Entscheidungen zu nehmen und nutzen dafür die Möglichkeiten politischer Kommunikation. Stabile politische Kommunikationsbeziehungen entstehen dann, wenn in ihnen Güter getauscht werden können, die der jeweils andere nicht hat, aber braucht (z. B. Publizität gegen Information oder Aufmerksamkeit gegen Nachrichten). Diese Akteurskonstellation hat sich in der letzten Zeit grundlegend gewandelt und wird von der Forschung sehr viel dynamischer gesehen als zuvor.

Die Dynamik drückt sich zum einen darin aus, dass die *tradierten Berufsrollen* der politischen Kommunikation ihre Ausschließlichkeit verlieren. An vielen Stellen finden sich in der Forschung Hinweise auf zusätzliche Kommunikatoren und veränderte Beziehungen. Für die verschiedenen Funktionen politischer Kommunikation haben sich differenzierte Berufsrollen herausgebildet. Im Zentrum steht nach wie vor der politische *Journalist* als der klassische „Schleusenwärter“ im Strom politischer Kommunikation. Dessen Berufsbild verändert sich unter dem Einfluss verschiedener Tendenzen, wie aus den kontinuierlichen Journalistenbefragungen hervorgeht (Weaver et al. 2006; Weischenberg et al. 2006). Zum einen verstärkt sich die Konkurrenz zwischen Anbietern, die im gleichen Medienbereich agieren, z. B. zwischen Rundfunkanbietern oder zwischen Tageszeitungen. Dies führt zu einer Verschärfung des publizistischen Wettlaufs um Informationen mit möglichst hohem Nachrichtenwert und um die Aufmerksamkeit der Rezipienten. Die Entwicklung kollidiert zusehends mit den tradierten Vorstellungen von journalistischer Qualität. Zum anderen verstärkt sich die Konkurrenz zwischen Anbietern aus verschiedenen Medienbereichen. So hat etwa die Abwanderung ganzer Anzeigenbereiche in Internetangebote die Ertragslage der Tageszeitungen stark geschwächt. Dadurch nimmt der Druck der Verlagsleitungen auf die Redaktionen, Kosten zu senken und zugleich medienübergreifend zu arbeiten, erheblich zu, da viele Verlage parallel ihr Engagement im Rundfunk und im Internet verstärken. Auch dies verändert die Normen, Abläufe und Organisationsformen journalistischer Arbeit. Prägnanter Ausdruck dessen ist die Einrichtung von „News-Desks“ in Redaktionen, an denen die Ressorts und Medienbereiche eines Anbieters koordiniert werden (Altmeppen 2006; Maier 2006). Ausdruck der Veränderung ist es auch, wenn vermehrt bei Online-medien oder Privatradios auf angelernte Kräfte statt auf ausgebildete Journalisten zurückgegriffen wird oder wenn auch in klassischen Medien verstärkt auf „user generated content“ gesetzt wird, also Laien als Kommunikatoren einbezogen werden (z. B. als „Leser-Reporter“ oder als Blogger; Ott 2006). Noch nicht hinreichend ist von der Forschung geklärt, welche publizistischen Konsequenzen die Veränderungen in der Anbieterstruktur nach sich ziehen, z. B. wenn Beteiligungsgesellschaften oder Telekommunikationsgesellschaften Anteile an Verlagen oder Rundfunksendern übernehmen.

Nicht allein aus Kostengründen hat sich der Einfluss von politischer *Public Relations* auf die Publikumsmedien verstärkt. Mehr noch ist dies darin begründet, dass Parteien, Verbände und staatliche Instanzen ihre PR-Aktivitäten intensiviert und professionalisiert haben (Tenscher 2003; Altmeyden et al. 2004; Negrine et al. 2007; allgemein zum Stand der PR-Forschung: Bentele et al. 2005). Der Kern der Professionalisierung ist darin zu sehen, dass PR-Verantwortliche die Medienlogik antizipieren und ihre Aktivitäten auf die Selektionsraster der Medien ausrichten – in zeitlicher Hinsicht (Berücksichtigung der täglichen, wöchentlichen und jährlichen Rhythmen der Medien), in sachlicher Hinsicht (Berücksichtigung von Nachrichtenfaktoren und Deutungsmustern der Medien) und in sozialer Hinsicht (Berücksichtigung von Leitmedien und von Redaktionsstrukturen). Die Professionalisierung schlägt sich in spezifischen Berufsbildern, Wissensbeständen und Ausbildungsgängen für PR im politischen Bereich nieder, oft unter dem Etikett „Public Affairs“. Einen Schritt weiter geht „Politisches Marketing“: Hinter dem Begriff steckt die Vermutung, dass sich die Programmatik politischer Organisationen stärker als früher an ihren jeweiligen Zielgruppen ausrichte, an Wählern, Unterstützern oder Sponsoren (Kreyher 2004). In welchem Maße sich „politisches Marketing“ nachweislich als programmatische Leitlinie in den Führungen politischer Organisationen durchgesetzt hat, ist umstritten. Studien machen deutlich, dass hier große Unterschiede zwischen Nationen und zwischen Organisationen zu finden sind (Lilleker/Lees-Marshment 2005; Davies/Newman 2006). Einen Sonderfall bilden NGOs, die Defizite im Hinblick auf andere Ressourcen dadurch kompensieren, dass sie die Öffentlichkeitsarbeit in das Zentrum ihrer Aktivitäten stellen (Röttger 2006; Voss 2007; Wimmer 2003).

Unübersichtlich ist die Entwicklung bei den Spezialisten, die politische Akteure im Hinblick auf politische Kommunikation *beraten*. So wächst das Heer an ehemaligen Politikern, die als Berater und Vermittler für diejenigen tätig werden, die Zugang zu politischen Entscheidern suchen. Zum anderen bieten vermehrt speziell für diesen Zweck gegründete Agenturen ihre Dienstleistungen an (Dagger et al. 2004). Es liegen erste empirische Ergebnisse über Arbeitsweise, Dienstleistungsspektrum und Klientenstruktur dieser Organisationen vor (Hoffmann et al. 2007).

Insgesamt zeigen sich in der Entwicklung der politischen Kommunikationsberufe gegenläufige Tendenzen: Zum einen gibt es Indizien für eine Professionalisierung der politischen Kommunikation im Sinne geregelter Ausbildung, theoretischer Fundierung und autonomer Regelung des Bereichs. Zum anderen findet man Hinweise auf eine De-Professionalisierung der politischen Kommunikation, am deutlichsten im Journalismus.

Die zweite grundlegende Veränderung der Akteurskonstellation besteht darin, dass nach Maßgabe der Forschung die *nationalen Grenzen* an Bedeutung für politische Kommunikation verlieren. Bislang postuliert das Akteursdreieck eine national begrenzte Öffentlichkeit, in der die Bürgerschaft eines Staates mit weitgehend innerhalb der staatlichen Grenzen operierenden Medienorganisationen und politischen Organisationen kommuniziert. Politische Kommunikation wurde und wird bevorzugt weitgehend im Rahmen nationaler Grenzen gedacht, die zumeist auch kulturelle Grenzen sind. Ein Großteil der empirischen Studien konzentriert sich auf Akteure, die innerhalb dieser Grenzen agieren, wie z. B. die Studie von Eilders et al. (2004) zum Zentrum der poli-

tischen Öffentlichkeit in Deutschland, dem „Kommentariat“. Auch die normativen Folien für politische Kommunikation sind für diesen Zuschnitt von Öffentlichkeit ausgelegt (Habermas 2006).

Dass diese Beschränkungen erodieren, sieht man daran, dass verstärkt nationale Kommunikationssysteme miteinander verglichen und auf ihre Leistungsfähigkeit hin geprüft werden (Hallin/Mancini 2004; Marcinkowski et al. 2006). Ebenso lässt sich dies mit Blick auf spezielle Aspekte wie die Professionalisierung politischer Kampagnenführung feststellen (Plasser/Plasser 2002). Der Blick *internationalisiert* sich: Man lässt ihn über die Grenzen schweifen und nach Kontrasten und Referenzen suchen.

Deutlich wird dies aber vor allem daran, dass die Forschung der Analyse *transnationaler* politischer Kommunikation sehr viel größere Aufmerksamkeit widmet als früher. Ein Teil der Arbeiten richtet sich auf die Frage, in welchem Maße von einer europäischen Öffentlichkeit die Rede sein kann. Die Antworten auf diese Frage fallen unterschiedlich aus. Nicht zuletzt hängen sie auch davon ab, ob sie aus politikwissenschaftlicher oder aus kommunikationswissenschaftlicher Perspektive gegeben werden (Hagen 2004; Peters et al. 2006).

Deutlich gewachsen ist das Interesse, das die Forschung der *Public Diplomacy* widmet (grundlegend dazu: Gilboa 2000; Melissen 2005). Die Karriere des Begriffs zeigt, wie wichtig für die (Außen-)Politik das kommunikative Einwirken auf (Teil-)Öffentlichkeiten anderer Nationen geworden ist. Mit Public Diplomacy ist der Anspruch verbunden, nicht nur Überzeugungsstrategien zu verfolgen, sondern auch die Verständigung zu fördern und den Dialog zu suchen. Darum wird immer auch interkulturelle Kommunikation als ein wesentlicher Bestandteil von Public Diplomacy genannt (Zöllner 2006). Die Nationen unterscheiden sich deutlich im Hinblick auf ihre Public Diplomacy. In Deutschland ist der zentrale mediale Akteur für auslandsorientierte Kommunikation die *Deutsche Welle* (Kleinsteuber 2002). In deren Programm wird das derzeit gültige Paradigma deutscher Kommunikations(außen)politik umgesetzt: Dialog der Kulturen, gerade auch mit der arabischen Welt. Hingegen markiert die Public Diplomacy der USA stärker die Unterschiede zwischen „westlicher“ und „muslimischer“ Welt (Lord 2006). Sie kann allerdings im Unterschied zur deutschen Public Diplomacy auf differenzierte Wirkungsanalysen zurückgreifen (z. B. El-Nawawy 2006; Kendrick/Fullerton 2004).

Der Blick in die Literatur zeigt, dass transnationale politische Kommunikation nicht allein von staatlichen oder staatsnahen Akteuren bestritten wird. Die Forschung hat sich besonders zwei weltweit operierenden Medienorganisationen zugewandt: Der arabische Sender *Al-Jazeera* wird zum einen deshalb eingehend gewürdigt, weil er eine eigenständige Quelle für Bildmaterial aus dem Nahen Osten darstellt und damit den einseitigen Nachrichtenfluss von Nord nach Süd durchbricht. Darüber hinaus wird er als Katalysator für demokratische Entwicklungen in der Region gesehen (Hafez 2004; Hahn 2005). Die besondere Bedeutung des anderen Anbieters wird dadurch deutlich, dass in der Literatur ein „CNN-Effect“ diskutiert wird (Robinson 2002). Hinter dieser Bezeichnung steckt die These, dass die bebilderte Berichterstattung aus Krisenregionen durch einen globalen Anbieter wie CNN (oder die besonders angesehene *BBC World*) einen Handlungsdruck auf politische Akteure erzeuge, dem sich diese nicht entziehen könnten. Durch die Thematisierung eines bestimmten Konflikts wird ein militärisches Eingreifen oder auch dessen Beendigung wahrscheinlicher, wie am Fall Somalia offen-

sichtlich wurde. In neueren Publikationen wird für eine differenzierte Betrachtung des Einflusses globaler Medienkommunikation auf die Prioritäten internationaler Politik plädiert (Gilboa 2005).

Dies verweist auf die Herausforderungen für die Akteurskonstellation, die von der veränderten Rolle der Medien im Krieg ausgehen. In den vergangenen Jahren ist die Forschung dazu sprunghaft angestiegen, was sich in thematischen Schwerpunkten von Fachzeitschriften (z. B. Eilders/Hagen 2005) und in etlichen Überblicksdarstellungen (z. B. Löffelholz 2004; McQuail 2006) niedergeschlagen hat. Im Mittelpunkt stehen dabei die Militäraktionen der USA seit dem 11. September, insbesondere der dritte Golfkrieg. Diese Kriege ließen die Veränderungen des Ensembles auf der internationalen Bühne deutlich hervortreten. Besondere Aufmerksamkeit hat die Steuerung von Information und Kommunikation durch die US-Regierung und das US-Militär hervorgerufen (Elter 2005). Beispiele für Einzelstudien sind die Darstellung der Tätigkeit des „Office for Strategic Influence“ im Afghanistan-Krieg (Christiansen 2004) und die Analyse der medialen Wahrnehmung von „Embedded Journalists“ im Irak-Krieg (Donsbach et al. 2005). In einem Teil der Literatur wird pointiert kritisiert, dass die US-Regierung in ihrem Kommunikationsmanagement zum Mittel der gezielten Desinformation gegriffen habe. Es ist sogar von einem „Revival of the Propaganda State“ die Rede (Snow/Taylor 2006; Miller 2004; Bussemer 2004). Die US-Erfahrungen zeigen, wie relevant ein professionelles Kommunikationsmanagement im Krieg geworden ist und wie schwierig: In einer vernetzten Welt können auch Kleingruppen oder Einzelpersonen mithilfe von Digitalkamera, Fotohandy und Internet als Kommunikatoren auftreten und gewinnen damit als Akteurstyp deutlich an Gewicht (Szukala 2005). Insbesondere terroristische Gruppen nutzen die Möglichkeiten der digitalen Medienwelt, um die Weltöffentlichkeit auf ihre Ziele aufmerksam zu machen, zumal terroristische Aktionen – etwa Exekutionen von Geiseln – aufgrund ihres Nachrichtenwerts auf große Publikumsresonanz stoßen: Terrorismus als Kommunikationsstrategie (Elter 2006; Neidhardt 2006).

Insgesamt macht die Forschung deutlich, in welchem Maße sich die Akteurskonstellation der politischen Kommunikation pluralisiert und globalisiert hat. Es werden sehr viel mehr und vor allem sehr viel mehr unterschiedliche Akteure in die Medienkommunikation einbezogen. Dadurch bieten sich mehr und andere Möglichkeiten der Koalition und der Konfrontation. Mit diesen Veränderungen korrespondieren methodische Innovationen in der Forschung. Ein Beispiel ist das „Words as Data“-Verfahren, mit dem die Positionen politischer Akteure präziser als bisher rekonstruiert werden können (Laver et al. 2003).

3. Veränderungen der politischen Kommunikation unter dem Inhaltsaspekt

Aus der Interaktion der Akteure im nationalen und internationalen Rahmen ergibt sich der politisch relevante Teil des Medienangebots. Analysen der Inhalte und der Formen zeigen große Unterschiede zwischen Medien und zwischen Nationen, z. B. im Hinblick auf den Anteil der Politik an der Berichterstattung, die Gewichtung von Themen, die politische Positionierung, die Berücksichtigung von Akteuren, die sprachliche Gestal-

tung oder den Grad an Visualisierung (z. B. de Vreese et al. 2006). Da diese intermedialen und internationalen Unterschiede nicht starr bleiben, gilt es auch Veränderungen im Medienangebot zu beschreiben und zu erklären. In der öffentlichen und fachöffentlichen Diskussion werden Tendenzen im politischen Medienangebot mit verschiedenen Schlagwörtern gekennzeichnet: Negativismus, Alarmismus und Beschleunigung (Politik werde heute schlechter, schriller und hektischer dargestellt als früher), Visualisierung und Personalisierung (Politik werde auf die bildlich darstellbaren und personell zurechenbaren Elemente reduziert), Individualisierung (politische Information werde von einem Gemeinschaftsgut zu einem Angebot für spezifische Zielgruppen) und vor allem Entpolitisierung und Entertainisierung (Politik nehme an Bedeutung im Medienangebot ab und werde, wenn überhaupt, dann nur noch von ihrer unterhaltsamen Seite dargestellt). Aus Politikberichterstattung sei „*Politainment*“ geworden (Dörner 2001). Die empirische Basis dieser Tendenzaussagen sind zumeist einzelne Beobachtungen, vor allem aus Wahlkampfzeiten, wo Home-Stories von Politikern und „Horse-Race“-Journalismus an die Stelle programmatischer Debatten getreten seien (Farnsworth/Lichter 2003). Diese Diagnosen sind in ihrer Pauschalität keineswegs durch systematische empirische Forschung gedeckt. Kontinuierliche Analysen der zeitlichen Verteilung von Programmsparten im Fernsehprogramm belegen vielmehr zumindest für Deutschland eine weitgehende Konstanz von Politikangeboten in den letzten Jahren. Dabei ist der Anteil der Politikberichterstattung am Gesamtprogramm bei den öffentlich-rechtlichen Sendern höher als bei der privaten Konkurrenz (ALM 2006; Krüger/Zapf-Schramm 2006 – mit sehr unterschiedlichen Operationalisierungen des Konzepts „politische Information“). Das Bild differenziert sich aber, wenn einzelne Formate genauer untersucht werden: So zeigen Inhaltsanalysen von Hauptnachrichtensendungen im Längsschnitt, dass in diesem Kernbereich der Informationsvermittlung der Politikteil tatsächlich zunehmend weniger umfassend ist und die Berichterstattung auch in öffentlich-rechtlichen Formaten emotionaler und personalisierter wird (Donsbach/Büttner 2005). Andererseits sind z. B. politische Talkshows wie „Sabine Christiansen“ besser als ihr Ruf. So wurde inhaltsanalytisch geprüft, ob die Auseinandersetzungen eher von sachorientierten Diskursen oder von den Selbstdarstellungen der Personen geprägt sind (Schultz 2006). Es zeigen sich signifikante Varianzen zwischen den verschiedenen Talkshows, aber auch zwischen verschiedenen Ausgaben einer bestimmten Show. Entscheidende Variablen sind z. B. die Zusammensetzung der Gesprächsrunde, die Klarheit des Themenzuschnitts und der Moderationsstil. Das Format „politische Talkshow“ bietet also Spielraum. Die Forschung hat sich mittlerweile auch der politischen Bezüge von fiktionalen Medienangeboten wie Spielfilmen oder Fernsehserien angenommen. Beispielhaft ist die Analyse der US-Fernsehserie „West Wing“, in der Arbeit und Alltag eines fiktiven amerikanischen Präsidenten und seines Mitarbeiterstabs dargestellt werden (Holbert et al. 2003). Beispielhaft ist diese Studie auch deshalb, weil sie über eine Inhaltsanalyse hinausging und durch eine Befragung von Nutzern zeigte, dass die Rezeption dieses Formats zu einer besseren Bewertung der US-Präsidentschaft insgesamt sowie von realen Präsidentenpersönlichkeiten führte.

Die Beispiele machen deutlich: Traditionell hat sich die Forschung zur politischen Kommunikation denjenigen Inhalten von Medien zugewandt, die bei den systematischen Inhaltsanalysen in unterschiedlicher Operationalisierung als „politische Informa-

tion “ gekennzeichnet wurden. Andere Teile des Medienangebots mit politischem Bezug hat man entweder ignoriert oder pauschal kritisiert. Dies ändert sich: Es werden auch diejenigen Inhalte politischer Kommunikation von der Forschung erfasst, die politische Bezüge aufweisen und auf Unterhaltungsbedürfnisse des Publikums abstellen. Man widmet ihnen gebührende Aufmerksamkeit – gebührend im Hinblick auf ihren Stellenwert im Medienangebot, gebührend im Hinblick auf die Motive der Zuwendung und gebührend im Hinblick auf Wirkungsvermutungen. Eine Grundlage für diese differenzierte Betrachtung des Verhältnisses von Politik und Unterhaltung bietet Saxer (2007). Er sieht die Entertainisierung medialer Politik als Teil einer grundlegenden Veränderung von Öffentlichkeit in der Mediengesellschaft. Normativ geprägte Theorien, wie z. B. die Theorie deliberativer Demokratie, seien nicht in der Lage, den rapiden Wandel der politischen Kommunikation aufzuklären und zu gestalten.

Auch eine andere kategorische Unterscheidung ist in der Forschung wesentlich durchlässiger geworden, die zwischen Wort und *Bild*. Traditionell hat sich die Forschung zur politischen Kommunikation auf den Wortanteil im politischen Medienangebot konzentriert. Ein Großteil der aktuellen Forschung setzt diese Tradition fort und analysiert z. B. die Berichterstattung zu kriegesischen Konflikten (Gleich 2003). So ergab die Analyse der Zeitungskommentare zum Afghanistan-Einsatz, dass – gemäß der Indexing-These (Bennett 2006) – die Medien ihre Positionen an die eindeutige Meinungsverteilung in der deutschen Politik anpassten (Pohr 2005). Aber die Forschung hat darüber hinaus die theoretische und methodische Herausforderung angenommen, die ikonischen Teile der politischen Kommunikation analytisch zu durchdringen – mit dem Ziel, die „Macht des Bildes“ (Frey 2005) zu fassen. Die Visualisierung des politischen Teils des Medienangebots ist kein neues Phänomen; sie hat aber an Tempo zugenommen, wie die Personalisierung der Wahlkämpfe (Knieper/Müller 2004), die zunehmende Bebilderung auch traditioneller Medien wie der Tageszeitung und die rasante Durchsetzung bildorientierter Medienangebote wie „YouTube“, „My Space“, „flickr“ oder demnächst „Joost“ deutlich indizieren. Wie die Forschung diesen „iconic turn“ aufgreift, soll wiederum an Beispielen zum Thema Medien und Krieg verdeutlicht werden: So wurde die Berichterstattung deutscher Fernsehsender zum Irak-Krieg daraufhin analysiert, ob Bilder verwundeter oder toter Soldaten gezeigt wurden (Weiß/Koch 2005). Ergebnis: Sie werden von den Spartenkanälen N-TV und N24 deutlich häufiger gezeigt als von den öffentlich-rechtlichen und privaten Vollprogrammen, weil die um größtmögliche Aktualität bemühten Nachrichtensender oft direkt und unkommentiert das Material ausländischer Anbieter übernehmen. Die Veränderung der Senderstruktur verändert auch die Inhalte und damit die Chancen, sie wahrzunehmen. Weiter vom Mainstream der traditionellen Inhalte und Formen politischer Kommunikation entfernt sich eine Untersuchung der Visualisierung von Krieg in Computerspielen (Klimmt et al. 2005). Kriegscomputerspiele können aus unterschiedlichen Perspektiven gespielt werden (z. B. First- versus Third-Person-Perspektive), sie bieten unterschiedliche Handlungsoptionen (z. B. planvolles Vorgehen versus schnelle Reaktion) und vielgestaltige Handlungsräume an (fiktive Szenarien versus reale militärische Konflikte als Vorbilder). Das Bild des Krieges ist dabei selektiv: Im Mittelpunkt stehen Kampfhandlungen, während der politische Kontext und die Folgen weitgehend ausgeblendet werden. Gerade dieses partielle Aufgreifen politischer Momente macht diesen Teil des Medienangebots für die Forschung zur politischen Kommunikation interessant.

Die Lösung aus der Fixierung an Information und Wort und die Hinwendung zu Unterhaltung und Bild sind nur unter bestimmten *methodischen Voraussetzungen* möglich. Die Forschung braucht Instrumente, mit denen Bildzeichen in angemessener Weise inhaltsanalytisch erfasst werden können. Die Erforschung visueller Kommunikation bedarf, so eine vielfach erhobene Forderung, einer methodologisch breiten Basis. Dafür sind Ansätze erkennbar, z. B. in der Systematisierung des Vorgehens bei der Bildanalyse (Müller 2003; Lester 2005) und in der Weiterentwicklung inhaltsanalytischer Verfahren für die Codierung nonverbaler Medieninhalte, insbesondere bewegter Bilder (Bente/Krämer 2004).

Zusammengefasst lässt sich sagen, dass die Inhalte politischer Kommunikation von der Forschung weiter gefasst werden als früher: Es werden vermehrt unterhaltungsorientierte Angebote mit einbezogen, sofern sie in einem politischen Kontext stehen. Zudem wird das Bild mehr als bisher gewürdigt – nicht mehr als Beigabe zum Text, sondern als eigenständiger „Text“. Dies geschieht beides auch mit Blick darauf, welche Inhalte und Formen vom Publikum wie intensiv rezipiert werden und welche Wirkung sie entfalten.²

4. *Veränderungen der politischen Kommunikation unter dem Aspekt der Rezeption und Wirkung*

Von der Forschung zur politischen Kommunikation wird in Wissenschaft und Öffentlichkeit vor allem erwartet, dass sie zuverlässige Antworten auf die vielen Fragen nach den Wirkungen des politischen Medienangebots findet. Zwei Problembereiche stehen im Mittelpunkt der aktuellen Forschung: Zum einen die folgenorientierte Frage, inwieweit die politischen Einstellungen durch Medien beeinflusst werden, insbesondere die für Wahlentscheidungen relevanten Einstellungen; und zum anderen die ursachenorientierte Frage, inwieweit die Ergänzung des Medienbouquets durch das Internet die politische Kommunikation verändert. Die Forschung hat in ihren Antworten eine Vielzahl spezieller theoretischer Ansätze entwickelt, die sich in unterschiedlichem Ausmaß empirisch bewährt haben. Fragen und Antworten kann man danach ordnen, auf welcher Ebene von politischer Kommunikation sie ansetzen. Die Kardinalfrage liegt auf der Mikroebene: In welchem Maße sind Veränderungen im politischen Denken, Wollen und Tun der einzelnen Bürger auf Veränderungen in den Medienangeboten und deren Nutzung zurückzuführen? Ergänzend finden sich Analysen, die auf der Mesoebene (Veränderungen von politischen Organisationen) oder auf der Makroebene (Veränderungen der Strukturen des politischen Systems) ansetzen.

² Bislang wenig ins Blickfeld der Forschung geraten ist der Wandel der Formen, der sich in der politischen Kommunikation vollzieht. Fernsehen, das über das Internet übertragen und empfangen wird, bietet ein völlig anderes Spektrum an Handlungsmöglichkeiten (Zeitsouveränität, selektiver Zugriff, Zusatzangebote, Parallelnutzung usw.) als das Fernsehen, das über Antenne oder Kabel übertragen und empfangen wird. Es sind folglich unterschiedliche Medien. Dies ist im Hinblick auf politische Kommunikation erst ansatzweise untersucht worden. Die Auseinandersetzung mit dem spezifisch *medienwissenschaftlichen* Forschungsprogramm, das auf die Materialität der Kommunikation oder, allgemeiner, auf die Medialität der Politik insgesamt abstellt, steht noch aus (Schicha 2003; Dörner 2006).

Ein großer Teil der Medienwirkungsforschung stellt auf die Beantwortung der Frage ab, ob und wie die Nutzung von Medienangeboten die individuellen *politischen Einstellungen* ändern oder stabilisieren kann. Dies können zum einen *themenspezifische* Einstellungen sein, wie z. B. die Haltung zum Euro (Maier et al. 2003). Die Forschung eröffnet damit auch neue Möglichkeiten für themenspezifische Kampagnen im Top-down-, Bottom-up- oder Media-centered-Modus (Kriesi 2004). Zum anderen können dies auch *generelle politische* Einstellungen sein, wie die Erwartungen an Politik (Iyengar/McGrady 2006) oder die Demokratiezufriedenheit (zur „Media-Malaise-These“: Maurer 2003; Wolling 2006). Insbesondere ist dabei die unter dem Inhaltsaspekt angeschnittene Frage von Belang, welchen Einfluss eine eher unterhaltungsorientierte Präsentation von Politik auf die Wahrnehmung von Politik generell hat und wie dies unter demokratiethoretischen Aspekten zu bewerten ist. Es wird z. B. gefragt, ob diese Formate eine politische Inklusionsfunktion erfüllen. Denn es werden mit ihnen Gruppen der Bürgerschaft erreicht, die sich anderen Medienangeboten mit politischem Zuschnitt entziehen (Baum 2002; Dohle et al. 2003). Noch weitreichender ist die Frage, inwieweit Medien Sozialkapital aufzehren und damit zur Erosion der Grundfesten von Gemeinschaft(en) beitragen (Hooghe 2002).

Im Zentrum der Forschung steht aber die Frage, welcher Einfluss auf die *wahlrelevanten Einstellungen* den Medien zuzurechnen ist. Im weit überwiegenden Teil der Forschung kommt man zu dem Schluss, dass der Einfluss der Medien erheblich und noch dazu gestiegen ist. Dies geht zulasten des Einflusses, der auf tradierte Überzeugungen zurückzuführen ist (Brettschneider 2005). Ob es auch zulasten des Einflusses unmittelbarer politischer Kommunikation geht, ist von vielen weiteren Faktoren abhängig (Schmitt-Beck 2004). Den Hintergrund dieses Befundes bildet die „Dealignment“-These (Dalton/Wattenberg 2000; Schoen/Weins 2005), die besagt, dass ein zunehmender Teil des Elektorats sich aus überkommenen (Partei-)Bindungen gelöst habe und dadurch offener für Medienbotschaften im unmittelbaren Vorfeld einer Wahl geworden sei (Kepplinger/Maurer 2005) – und zwar für Botschaften zum Verhalten der anderen Wähler (zum Einfluss veröffentlichter Demoskopie: Raupp 2007), zu den Kandidaten (zum Einfluss der Medien auf das Kandidatenbild: Maurer/Reinemann 2003; Klein/Rosar 2005) und zu den Sachfragen des Wahlkampfes.

Dieser letzte Aspekt hat in der Forschung besondere Aufmerksamkeit erfahren. Die gestiegene Volatilität der Wählerschaft eröffnet dem Themenmanagement im Wahlkampf neue Möglichkeiten. Dies beginnt beim *Agenda-Setting* – eine mittlerweile durch empirische Forschung breit abgesicherte Wirkung der Medien auf die Einstellungen der Wähler (McCombs 2004). Medien gewinnen demnach politischen Einfluss vor allem dadurch, dass sie vorgeben, *worüber* die Wähler nachdenken und mit anderen sprechen, weniger dadurch, dass sie vorgeben, *was* die Wähler denken und mit anderen besprechen. Diese Überlegung wurde in zwei Richtungen weiterentwickelt. Zum einen wurde über den „Priming“-Ansatz der Bogen zu den Wahlentscheidungen geschlagen: Wenn in den Medien ein Thema als besonders relevant vermittelt wird und wenn eine Partei bei diesem Thema als besonders kompetent angesehen wird, dann steigen die Wahlchancen dieser Partei (Scheufele/Tewsbury 2007). Eine zweite Entwicklung firmiert unter dem Etikett „Second-Level-Agenda-Setting“ (Rössler 2006). Dabei wird den Medien zugerechnet, dass sie den Rezipienten auch vermitteln, *wie* diese über die

prioritären Themen nachdenken sollen, indem in den Medien bestimmte Charakteristika der Themen hervorgehoben und in ein normativ gefärbtes Licht gestellt werden.

Durch diese Erweiterung nähert sich die Agenda-Setting-Forschung dem *Framing-Ansatz* an, der in den letzten Jahren stark an Bedeutung gewonnen hat. Ihm zufolge werden in Wahlkampagnen nicht nur Themen durchzusetzen versucht, sondern die Kampagnen zielen darauf, mithilfe von Medien Deutungsmuster zu vermitteln oder zumindest eine solche Vermittlung zu beeinflussen. Dies lässt sich an der Konfrontation der Thematiken „soziale Gerechtigkeit“ und „wirtschaftliche Effizienz“ im letzten Bundestagswahlkampf illustrieren. Um zu untersuchen, wie Frames entstehen und wie sie sich verändern, wird von der Forschung an ganz unterschiedlichen Stellen angesetzt: bei politischen Kommunikatoren, Journalisten, Rezipienten und Medieninhalten. Frames prägen sowohl die Entstehung und Verbreitung von Aussagen als auch ihre Rezeption und die Fähigkeit, sich an den Inhalt zu erinnern. Wegen dieser großen Bandbreite sind sehr unterschiedliche Fassungen des grundlegenden Konzepts in der Literatur zu finden (vgl. Scheufele 2003; Matthes/Kohring 2004; Dahinden 2006). Die Forschungsansätze haben aber einen gemeinsamen Kern: Frames werden als Interpretationsrahmen gesehen, die dazu dienen, Informationen einzuordnen und zu verarbeiten. Frames steigern dadurch die Effizienz von Kommunikationsprozessen. Mit Frames können nicht nur einzelne Elemente eines Themas herausgehoben und andere ignoriert werden; es können auch Ereignisse und deren Folgen in einen Kontext gestellt werden, der bestimmte Schlussfolgerungen nahelegt und andere nicht. Eine solche Rahmung von Ereignissen durch die Frames in den Medienbotschaften beeinflusst die Vorstellungen von Rezipienten bzw. Wählern, aber auch ihre Bewertungen, Schlussfolgerungen und damit letztlich auch ihre Entscheidungen.

Über den Framing-Ansatz wird ein Anschluss an die kognitionspsychologisch gut abgesicherte *Schematheorie* möglich (beispielhaft: Tiele/Scherer 2004). Informationsaufnahme und -verarbeitung hängen dieser Theorie zufolge von stabilen kognitiven Strukturen der Rezipienten ab. Sie nehmen mediale Inhalte in dem Maße auf, wie sie in bereits vorhandene Muster integriert werden können (Matthes 2004; siehe auch Schoen 2006). Mit diesem Brückenschlag zur Schematheorie werden Grundfragen der Medienwirkungsforschung aktualisiert: Inwieweit begrenzen kognitive Muster die Einflussmöglichkeiten von Medien? Inwieweit werden aber auch die kognitiven Muster langfristig durch Medieneinflüsse verändert?

Methodisch sind neben die Auswertung von Aggregat- und Individualdaten vermehrt Kombinationen von Befragung und Inhaltsanalyse getreten (z. B. Scheufele et al. 2005). Es finden sich auch mehr Anwendungsvarianten experimenteller Verfahren als früher. So wurde in Laborexperimenten untersucht, welche Wirkung die Kandidatenduelle vor der Bundestagswahl 2002 auf Zuschauer hatten (Faas/Maier 2004) oder ob unterschiedliche Zeitungsbilder eines Kandidaten einen Einfluss darauf haben, wie die Wähler den Kandidaten einschätzen und ob sie sich für oder gegen ihn entscheiden (Barrett/Barrington 2005). Zudem werden neben postrezeptiven auch rezeptionsbegleitende Daten erhoben und ausgewertet (Maurer/Reinemann 2003).

Auf der *Mesoebene* der Forschung zur Wahlkommunikation lautet die Leitfrage: Wie reagieren die politischen Akteure darauf, dass der Einfluss der Medien auf die Wahlentscheidungen gestiegen ist? Inwieweit sind also Veränderungen in politischen

Organisationen direkt oder indirekt auf Medieneinflüsse zurückzuführen? Dies schließt den Bogen zum eingangs behandelten Akteursaspekt. Zwar ist diese Wirkungsdimension wesentlich weniger gut erforscht, aber es gibt deutliche Hinweise darauf, dass sich politische Akteure stärker als früher auf die Erfordernisse der medialen Kommunikation einstellen und sich dies in ihrer Aufbau- und Ablauforganisation niederschlägt (Donges 2006). Dies betrifft nicht nur die Kommunikation nach außen. Vor allem die organisationsinterne Kommunikation hat sich durch internetbasierte Medien gewandelt. Ihre Nutzung ermöglicht eine schnellere und effizientere Koordination von Basis und Spitze (z. B. als unmittelbare Reaktion einer Partei auf Aktionen des politischen Gegners); es ist aber auch umgekehrt eine schnellere und effizientere Koordination von Teilen der Basis gegen die Organisationsspitze möglich (Kamps/Nieland 2006).

Auf der *Makroebene* ist die Leitfrage: In welchem Maße sind grundlegende Veränderungen im politischen System auf Veränderungen in der medialen Kommunikation zurückzuführen? Als Antwort darauf sind weit reichende Thesen formuliert worden, z. B. dass die „Mediendemokratie“ die Parteiendemokratie abgelöst habe oder dass sich ein neuer Herrschaftstyp der „Mediokratie“ (Meyer 2001) oder der „Videocracy“ (Mazzeni 1995) herausgebildet habe. Weniger spekulative Antworten kann die komparative Forschung zur politischen Kommunikation in verschiedenen Staaten geben, z. B. zur Veränderung von Wahlkampagnen (Farrell/Schmitt-Beck 2006; Wagner 2005). Auch die Möglichkeiten, die die Europawahlen für den Systemvergleich eröffnen, werden mittlerweile genutzt (de Vreese et al. 2006; Holtz-Bacha 2005; Peter et al. 2004; Tenschner 2005). Die internationalen Vergleiche beschränken sich nicht auf Wahlkommunikation. So ist unter dem Stichwort „Reformkommunikation“ ausgelotet worden, welche Strategien in welchem politisch-kulturellen Umfeld zielführend sind und welche nicht (Korte 2007). Noch weiter gehen Studien, die politische Kommunikationskulturen konfrontieren (Pfetsch 2003) oder den Zusammenhang von Mediensystem, Governance-Qualität und Lebensqualität vergleichend untersuchen (Norris 2004).

In einem anderen Strang der Forschung werden ebenfalls neue Wirkungspotenziale sichtbar: Die Erweiterung des medialen Spektrums durch das *Internet* erlaubt auch in politischer Hinsicht eine Kommunikation, in der die räumlichen, zeitlichen und sozialen Grenzen herkömmlicher Massenkommunikation obsolet geworden sind. Diese Möglichkeiten werden in der politischen Diskussion unterschiedlich eingeschätzt und ambivalent bewertet. Das ist nicht überraschend: Wann immer sich neue Medien durchsetzen, verbanden sich damit Hoffnungen und Befürchtungen mit Blick auf Veränderungen der politischen Kommunikation. Den spekulativen Prognosen hat die Forschung einige belastbare Antworten auf die Frage entgegensetzen, welche Veränderungen der politischen Kommunikation mit dem rapiden Siegeszug des Internets einhergehen.

Auf der *Mikroebene*, also im Hinblick auf die individuelle politische Kommunikation, haben verschiedene Panel-Studien mit unterschiedlichem methodischen Design erbracht, dass diejenigen, die sich einen Internetanschluss zulegen, auch in politischer Hinsicht ihr Informationsverhalten verändern. Zumeist zeigt sich ein Mobilisierungseffekt im Hinblick auf verschiedene politische Aktivitäten (Bimber 2003; Emmer 2005; Johnson/Kaye 2003; Tolbert/McNeal 2003). Die Veränderungen sind längst nicht so eklatant wie in anderen Lebensbereichen, z. B. beim Konsumverhalten, aber nachweisbar, auch wenn andere Faktoren kontrolliert werden (Hardy/Scheufele 2006).

Die deskriptive Forschung hat gezeigt, dass die Veränderungen in spezifischen Gruppen stärker ausgeprägt sind und sich sehr rasch vollziehen. So zeigen kontinuierliche Befragungen von Jugendlichen, wie sehr mittlerweile Handy, Computer und Internet die Medienwelt dominieren. Von den Informationsangeboten der öffentlich-rechtlichen Sender und der Printmedien wird diese Gruppe nur noch auf Umwegen erreicht (Medienpädagogischer Forschungsverband Südwest 2006). In den Analysen von Umfragedaten wird über den engen Kreis der Jugendlichen hinaus eine Gruppe erkennbar, die traditionelle Formen der politischen Kommunikation meidet, aber internetbasierte Medien auch für politische Zwecke intensiv nutzt (Emmer et al. 2006). Zwar haben sich die Befürchtungen, es werde durch das Internet zu einer dauerhaften Spaltung der Bevölkerung in Nutzer und Nicht-Nutzer des Internets kommen, als grundlos erwiesen; es zeigten sich jedoch prägnante Unterschiede im politischen Informationsverhalten nach Bildung und Alter (differenziert zu Befürchtungen und zu nachweisbaren Unterschieden: Marr 2005).

Etwas schwer tut sich die Forschung bislang noch mit den neuen internetbasierten Kommunikationsplattformen, an die sich die zweite Hochphase der Erwartungen an das Internet knüpft („Web 2.0“). Die neuen Angebote bieten durch einfach handhabbare und kostenlos erhältliche Instrumente auch politisch nutzbare Möglichkeiten, über die Rezipientenrolle hinaus Kommunikatorenfunktionen zu übernehmen („Broadcast Yourself“). Einer der Gründe für die Zurückhaltung der Forschung dürfte gerade darin liegen, dass die klassischen Rollenvorgaben der (politischen) Kommunikationsforschung in Form von „Publikum“, „Kommunikator“ oder „Rezipient“ auf diesen Teil der Medienwelt nicht mehr passen. Bislang fehlen sowohl systematische Inhaltsanalysen, etwa zum politischen Gehalt der Selbstpräsentationen auf diesen Plattformen, als auch methodisch kontrollierte Studien zu den Motiven, Qualitätsmaßstäben, Nutzungsgewohnheiten und sozialen Merkmalen der „netizens“. Andererseits liegen mittlerweile einige Studien zur wissenschaftlich kontrollierten Nutzung von Online-Foren zur Bürgerbeteiligung vor – auch in der Form von Feldexperimenten (Schweitzer 2004; Fishkin/Luskin 2005).

Auch im Hinblick auf das Internet sind die politischen Veränderungen auf der *Mesoebene* wesentlich weniger untersucht. Für die generellen Verschiebungen in der Anbieterstruktur durch die Digitalisierung der Medienprodukte liegen empirisch gut abgesicherte Modelle vor (Anderson 2006). Erkennbar ist auch, dass etablierte politische Akteure und Medien auf die Verschiebungen in der individuellen politischen Kommunikation mit teilweise beachtlichem Erfolg reagieren. Noch auf der Ebene von Fallstudien und praktischen Anleitungen ist die Literatur zum systematischen Einsatz der digitalen Medien in politischen Kampagnen für die punktgenaue und iterative Ansprache von kleinen Wählergruppen und einzelnen Wählern, wie etwa nach dem Vorbild des Marketings im Konsumgüterbereich (Merz et al. 2006). Die strategische politische Kommunikation verändert sich durch das Internet nicht nur deswegen, weil die Kosten für persönliche Ansprache und Dialog radikal sinken, sondern auch, weil die Mediatisierung der individuellen politischen Kommunikation für eine Fülle von Daten sorgt, aus denen sich kommunikative Profile bilden lassen, welche wiederum als Grundlage für die weitere Politikvermittlung dienen können. Es liegen keine belastbaren Forschungsergebnisse darüber vor, in welchem Ausmaß etablierte politische Akteure

diese Chancen nutzen. Hingegen hat sich die Forschung durchaus damit beschäftigt, welche neuen Konstellationen sich mit neuen „Gate-Keepern“ ergeben. So nehmen „Google“ – als Beispiel für einen kommerziellen Informationsdienstleister – oder „Wikipedia“ – als Beispiel für eine kollaborative Organisationsform – Schlüsselstellungen im politischen Informationsprozess ein (Machill/Beiler 2007; Möller 2006). Auch netzgestützte Dienstleistungen für politische Kommunikation, wie Orientierungshilfen bei Wahlentscheidungen und Volksabstimmungen, gewinnen an Bedeutung. Die begleitende Forschung zeigt, dass diese Angebote vorzugsweise von einer Informationselite genutzt werden (Marschall 2005; Schmitt-Beck et al. 2005).

Die Forschung ist zudem daran interessiert, herauszufinden, welche Veränderungen damit auf der *Makroebene* politischer Kommunikation verbunden sind. So wird z. B. eine Fragmentierung der politischen Öffentlichkeit wahrgenommen (Fohrmann/Orzessek 2002; Imhof 2003). In der Debatte darüber ist schwer einzuschätzen, wie viel der konstatierten Veränderung darauf zurückzuführen ist, dass sich die politische Kommunikation selbst verändert hat, und wie viel darauf, dass sich die Wahrnehmung der Veränderung verändert hat. Denn auch die Analyseinstrumente und die Zugangsmöglichkeiten haben sich verbessert. Zwar wurden auch zu früheren Zeiten milieuspezifische Teile der Öffentlichkeit untersucht, aber zu vielen „Stammtischen“ einzelner sozialer und politischer Gruppen bekam die Forschung keinen Zugang, weil die Kommunikation abgeschirmt und flüchtig war. Folglich blieb dieser Teil der politischen Kommunikation weitgehend unbeobachtet. Heute ist der „Stammtisch“ dadurch, dass er im Internet steht, hörbarer und sichtbarer geworden und kann deshalb besser dokumentiert und untersucht werden.

Zusammenfassend ist festzuhalten, dass die neuen Wirkungspotentiale die Forschung zur politischen Kommunikation vor Herausforderungen stellen, welche sie auf der Mikroebene angenommen hat – insbesondere durch die Adaption psychologischer und soziologischer Methodik in Datenerhebung und Datenanalyse. In Zukunft ist zu erwarten, dass verstärkt auf die rezeptionsorientierte Unterhaltungsforschung (Früh 2002; Bryant/Vorderer 2006; Wirth et al. 2006) und auf emotionspsychologische Ansätze zurückgegriffen werden wird, um auch affektive Wirkungen bei der Rezeption berücksichtigen zu können (vgl. Wirth/Schramm 2005). Die Veränderungen auf der Meso- und Makroebene stellen die Forschung zur politischen Kommunikation – zumindest im Hinblick auf methodische Innovationen – vor noch größere Herausforderungen.

5. Fazit

Politische Kommunikation ist ein Feld im Umbruch. Viele der herkömmlichen Begrenzungen und Einteilungen sind hinfällig geworden, neue noch nicht immer klar erkennbar. Unter allen drei Aspekten wurde deutlich, wie tiefgreifend der Wandel ist:

Die *Akteurskonstellation* der politischen Kommunikation erweitert sich in zwei Richtungen: in Richtung Pluralisierung und Globalisierung. Die Kommunikation klassischer politischer Akteure nach eingespielten Regeln in nationalen Grenzen behält ihre Bedeutung, aber sie ändert ihre Gestalt dadurch, dass in einem weltweiten Horizont

der Wahrnehmung kommuniziert wird und dass jederzeit neue Akteure hinzutreten können, wie der Streit um die Mohammed-Karikaturen anschaulich gezeigt hat.

Bei den *Inhalten* der politischen Kommunikation zeigen sich zwei weitere Tendenzen: Zwar behalten klassische Inhalte ihre Bedeutung, aber sie werden ergänzt durch visuelle Elemente und durch Inhalte, die stärker auf Unterhaltung orientiert sind – also Spannung, Überraschung und Amüsement mit sich bringen –, ohne dabei den politischen Kontext zu verlassen.

Individualisierung und Digitalisierung sind die Tendenzen, von denen die Entwicklung der politischen Kommunikation unter dem *Wirkungsaspekt* gekennzeichnet ist. Die sozialen Voraussetzungen für politische Kommunikation haben sich verändert, wie sich besonders deutlich in Wahlkampagnen zeigt. Dadurch wachsen die Einflussmöglichkeiten über politische Kommunikation. Das Spektrum der internetbasierten Medien erlaubt es, effizient diese Möglichkeiten zu nutzen.

Alle diese Tendenzen verstärken einander – mit dem Ergebnis einer Beschleunigung des Wandels politischer Kommunikation. Der Wandel ist unmittelbar erfahrbar geworden und wird seinerseits zu einem Thema der Medien – Medienberichterstattung über die Medienberichterstattung (Esser et al. 2005; Weiß 2005).

Die *Forschung* zur politischen Kommunikation hat diesen Umbruch angenommen und ihre thematischen Prioritäten, theoretischen Ansätze und methodischen Designs überprüft. Sie hat über angewandte Forschung diesen Wandel noch intensiviert. Die Herausforderungen anzunehmen, war ihr vor allem deshalb möglich, weil die Forschung in einem Bereich stattfindet, in dem sich disziplinäre Sichtachsen kreuzen. Die Überschneidung politik- und kommunikationswissenschaftlicher Perspektiven ist produktiv, wie die Herausbildung fachspezifischer Foren (Fachzeitschriften, Buchreihen) und gelungene Institutionalisierungen (Ausbildungsgänge, Fachgesellschaften) zeigen. Sie ist innovativ, weil hier unterschiedliche Wissenschaftskulturen aufeinander treffen, die sich jeweils aus ihrer Binnengliederung lösen müssen: „Politische Kommunikation“ ist kein Konzept, das allein den internationalen Beziehungen, der politischen Theorie oder dem Systemvergleich zuzurechnen wäre; und es ist kein Konzept, das allein der Kommunikator-, der Angebots- oder der Wirkungsforschung zugewiesen werden könnte. Dadurch ergibt sich jeweils die Chance einer übergreifenden Sichtweise. Fruchtbar ist das Zusammentreffen unterschiedlicher Wissenschaften auch deshalb, weil Politik- und Kommunikationswissenschaft jeweils Einflüsse weiterer Wissenschaften einbringen. So gelangen über die Politikwissenschaft ökonomische Konzepte in die Analyse der Akteurskonstellation politischer Kommunikation und über die Kommunikationswissenschaft psychologische Konzepte in die Analyse der politischen Medienwirkungen. Insofern ist der beschleunigte Wandel des Gegenstandsbereichs nicht die einzige Quelle für die Forschungsentwicklung: Auch der Austausch zwischen den Wissenschaften und die fachimmanente Verknüpfung von Theorien und Methoden sind Quellen der Innovation.

Die Herausforderung durch den rapiden Wandel der politischen Kommunikation anzunehmen, birgt auch Probleme: Der Anspruch auf Aktualität erfordert es, mit dem Wandel des Gegenstandsbereichs Schritt zu halten. Entsprechend selten sind historische Rückblicke und Längsschnittanalysen, die es erfordern, einen Schritt vom aktuellen Geschehen und einzelnen Fällen zurückzutreten, um die Dynamik insgesamt erfassen

zu können. Nur dadurch wäre es aber möglich, die unübersichtliche Entwicklung in theoretisch basierten Konzepten zu modellieren. Noch ist nicht hinreichend geklärt, ob die Forschung zur politischen Kommunikation einen eigenständigen Beitrag leistet, um relevante politische Phänomene zu erklären, z. B., worin sich soziale Sicherungssysteme unterscheiden oder warum Konflikte in Kriege übergehen. Die Perspektive der politischen Kommunikation ermöglicht für einen Teil dieser Phänomene belastbare Deutungen und Erklärungen, an denen ein aufgeklärtes politisches und mediales Handeln anschließen kann. „Kommunikation“ ist damit zu einem Schlüsselbegriff für eine unter mehreren wissenschaftlichen Sichtweisen geworden. Andere Wissenschaftsperspektiven versuchen, relevante politische Phänomene anders zu erklären, z. B. aus den Erfordernissen exogener Faktoren oder aus der Herrschaftsstruktur. Medien spielen bei diesen Ansätzen eine untergeordnete Rolle. Die Vielfalt möglicher Perspektiven ist eine Voraussetzung fruchtbarer Forschung. Empirische Forschung ermöglicht aber auch ein intersubjektiv gültiges Urteil darüber, welche Perspektive zu einem bestimmten Zeitpunkt eine klarere Sicht auf die Welt und ein umsichtigeres Handeln zulässt.

Literatur

- ALM (Hg.), 2005: ALM Programmbericht. Fernsehen in Deutschland 2005. Berlin.
- Altmeppen, Klaus-Dieter, 2006: Journalismus und Medien als Organisation. Leistungen, Strukturen und Management. Wiesbaden.
- Altmeppen, Klaus-Dieter/Röttger, Ulrike/Bentele, Günter (Hg.), 2004: Schwierige Verhältnisse. Interdependenzen zwischen Journalismus und PR. Wiesbaden.
- Anderson, Chris, 2006: The Long Tail. New York.
- Barrett, Andrew W./Barrington, Lowell W., 2005: Is a Picture Worth a Thousand Words? Newspaper Photographs and Voter Evaluations of Political Candidates, in: The Harvard International Journal of Press/Politics 10, 98-113.
- Baum, Matthew A., 2002: Sex, Lies, and War: How Soft News Brings Foreign Policy to the Inattentive Public, in: American Political Science Review 96, 91-109.
- Bennett, W. Lance, 2006: News: The Politics of Illusion. New York.
- Bente, Gary/Krämer, Nicole C., 2004: Inhaltsanalyse medialer Angebote, in: Mangold, Roland/Vorderer, Peter/Bente, Gary (Hg.), Lehrbuch der Medienpsychologie. Göttingen, 201-229.
- Bentele, Günter/Fröhlich, Romy/Szyszk, Peter (Hg.), 2005: Handbuch der Public Relations. Wissenschaftliche Grundlagen und berufliches Handeln. Wiesbaden.
- Bimber, Bruce, 2003: Information and American Democracy. Technology in the Evolution of Political Power. Cambridge u. a.
- Bimber, Bruce/Davis, Richard, 2003: Campaigning Online. The Internet in U. S. Elections. Oxford/New York.
- Brettschneider, Frank, 2005: Massenmedien und Wählerverhalten, in: Falter, Jürgen/Schoen, Harald (Hg.), Handbuch Wahlforschung. Wiesbaden, 473-500.
- Bryant, Jennings/Vorderer, Peter (Hg.), 2006: Psychology of Entertainment. Mahwah.
- Bussemer, Thymian, 2004: Medien als Kriegswaffe. Eine Analyse der amerikanischen Militärpropaganda im Irak-Krieg, in: Aus Politik und Zeitgeschichte B 49/50, 20-28.
- Christiansen, Olaf, 2004: Das „Office for Strategic Influence“ als Informationsmanager in der US-Kampagne in Afghanistan, in: Publizistik 49, 66-81.
- Dagger, Steffen/Greiner, Christoph/Leinert, Kirsten/Melß, Nadine/Menzel, Anne (Hg.): 2004: Politikberatung in Deutschland. Praxis und Perspektiven. Wiesbaden.
- Dahinden, Urs, 2006: Framing. Eine integrative Theorie der Massenkommunikation. Konstanz.
- Dalton, Russell J./Wattenberg, Martin P. (Hg.), 2000: Parties without Partisans: Political Change in Advanced Industrial Democracies. Oxford.

- Davies, John P./Newman, Bruce I. (Hg.), 2006: *Winning Elections with Political Marketing*. Binghampton.
- De Vreese, Claes H./Banducci, Susan A./Semetko, Holli A./Boomgaarden, Hajo G., 2006: The News Coverage of the 2004 European Parliamentary Election Campaign in 25 Countries, in: *European Union Politics* 7, 477-504.
- Dörner, Andreas, 2001: Politainment. Politik in der medialen Erlebnisgesellschaft. Frankfurt a. M.
- Dörner, Andreas, 2006: Medienkultur und politische Öffentlichkeit: Perspektiven und Probleme der Cultural Studies aus politikwissenschaftlicher Sicht, in: Hepp, Andreas/Winter, Rainer (Hg.), *Kultur – Medien – Macht*. Wiesbaden, 219-236.
- Dohle, Marco/Wirth, Werner/Vorderer, Peter, 2003: Emotionalisierte Aufklärung. Eine empirische Untersuchung zur Wirkung der Fernsehserie „Holokaust“ auf antisemitisch geprägte Einstellungen, in: *Publizistik* 48, 288-309.
- Donges, Patrick, 2006: Medien als Institution und ihre Auswirkungen auf Organisationen. Perspektiven des soziologischen Neo-Institutionalismus für die Kommunikationswissenschaft, in: *Medien & Kommunikationswissenschaft* 54, 563-578.
- Donges, Patrick (Hg.), 2007: *Von der Medienpolitik zur Media Governance?* Köln.
- Donsbach, Wolfgang/Büttner, Katrin, 2005: Boulevardisierungstrend in deutschen Fernsehnachrichten. Darstellungsmerkmale der Politikberichterstattung vor den Bundestagswahlen 1983, 1990 und 1998, in: *Publizistik* 50, 21-38.
- Donsbach, Wolfgang/Jandura, Olaf/Müller, Diana, 2005: Kriegsberichterstattung oder willfährige Propagandisten? Wie deutsche und amerikanische Printmedien die „Embedded Journalists“ im Irak-Krieg sahen, in: *Medien & Kommunikationswissenschaft* 53, 298-313.
- Eilders, Christiane/Hagen, Lutz, 2005: Kriegsberichterstattung als Thema kommunikationswissenschaftlicher Forschung. Ein Überblick zum Forschungsstand und den Beiträgen in diesem Themenheft, in: *Medien & Kommunikationswissenschaft* 53, 205-221.
- Eilders, Christiane/Neidhardt, Friedhelm/Pfetsch, Barbara, 2004: *Die Stimme der Medien. Pressekommentare und politische Öffentlichkeit in der Bundesrepublik*. Wiesbaden.
- El-Nawawy, Mohammed, 2006: US Public Diplomacy in the Arab World: The News Credibility of Radio Sawa and Television Alhurra, in: *Global Media and Communication* 2, 183-203.
- Elter, Andreas, 2005: *Die Kriegsverkäufer. Geschichte der US-Propaganda 1917-2005*. Frankfurt a. M.
- Elter, Andreas, 2006: Die RAF und die Medien. Ein Fallbeispiel für terroristische Kommunikation, in: Kraushaar, Wolfgang (Hg.), *Die RAF und der linke Terrorismus*. Band 2. Hamburg, 1060-1074.
- Emmer, Martin, 2005: Politische Mobilisierung durch das Internet? Eine kommunikationswissenschaftliche Untersuchung zur Wirkung eines neuen Mediums. München.
- Emmer, Martin/Füting, Angelika/Vowe, Gerhard, 2006: Wer kommuniziert wie über politische Themen? Eine empirisch basierte Typologie individueller politischer Kommunikation, in: *Medien & Kommunikationswissenschaft* 54, 216-236.
- Esser, Frank/Schwabe, Christiane/Wilke, Jürgen, 2005: Metaberichterstattung im Krieg. Wie Tageszeitungen die Rolle der Nachrichtenmedien und der Militär-PR in den Irakkonflikten 1991 und 2003 framen, in: *Medien & Kommunikationswissenschaft* 53, 314-332.
- Faas, Thorsten/Maier, Jürgen, 2004: Mobilisierung, Verstärkung, Konversion? Ergebnisse eines Experiments zur Wahrnehmung der Fernsehduelle im Vorfeld der Bundestagswahl 2002, in: *Politische Vierteljahresschrift* 45, 55-72.
- Farnsworth, Stephen J./Lichter, Robert, 2003: *The Nightly Nightmare: Network Television's Coverage of U.S. Presidential Elections, 1988-2000*. Lanham.
- Farrell, David/Schmitt-Beck, Rüdiger (Hg.), 2006: *Do Political Campaigns Matter? Campaign Effects in Elections and Referendums*. London/New York.
- Fishkin, James S./Luskin, Robert C., 2005: Experimenting with a Democratic Ideal: Deliberative Polling and Public Opinion, in: *Acta Politica* 40, 284-298.
- Fohrmann, Jürgen/Orzessek, Arno (Hg.), 2002: *Zerstreute Öffentlichkeiten. Zur Programmierung des Gemeinnsinns*. München.

- Frey, Siegfried, 2005: Die Macht des Bildes. Der Einfluss der nonverbalen Kommunikation auf Kultur und Politik. Bern.
- Früh, Werner, 2002: Unterhaltung durch das Fernsehen. Eine molare Theorie. Konstanz.
- Gilboa, Eytan, 2000: Mass Communication and Diplomacy: A Theoretical Framework, in: Communication Theory 10, 275-309.
- Gilboa, Eytan, 2005: The CNN Effect: The Search for a Communication Theory in International Relations, in: Political Communication 22, 27-44.
- Gleich, Uli, 2003: Qualität im Journalismus am Beispiel der Kriegsberichterstattung. Forschungsbeiträge zur Qualitätsdebatte, in: Media Perspektiven o. Jg., 139-148.
- Graber, Doris A., 2005: Political Communications Faces the 21st Century, in: Journal of Communication 55, 479-507.
- Habermas, Jürgen, 2006: Political Communication in Media Society: Does Democracy still Enjoy an Epistemic Dimension? The Impact of Normative Theory on Empirical Research, in: Communication Theory 16, 411-426.
- Hafez, Kai, 2004: Arabisches Satellitenfernsehen – Demokratisierung ohne politische Parteien?, in: Aus Politik und Zeitgeschichte B 48, 17-23.
- Hagen, Lutz (Hg.), 2004: Europäische Union und mediale Öffentlichkeit. Theoretische Perspektiven und empirische Befunde zur Rolle der Medien im europäischen Einigungsprozess. Köln.
- Hahn, Oliver, 2005: Arabisches Satelliten-Nachrichtenfernsehen. Entwicklungsgeschichte, Strukturen und Folgen für die Konfliktberichterstattung aus dem Nahen und Mittleren Osten, in: Medien & Kommunikationswissenschaft 53, 241-260.
- Hallin, Daniel C./Mancini, Paolo, 2004: Comparing Media Systems. Three Models of Media and Politics. Cambridge.
- Hardy, Bruce W./Scheufele, Dietram A., (2006): The Internet and Democratic Citizenship, in: Antti-roiko, Ari-Veikko/Mälkiä, Matti (Hg.), Encyclopedia of Digital Government. Hershey, 1250-1254.
- Hoffmann, Jochen/Steiner, Adrian/Jarren, Otfried, 2007: Politische Kommunikation als Dienstleistung. Public-Affairs-Berater in der Schweiz. Konstanz.
- Holbert, R. Lance/Pillion, Owen/Tschida, David A./Armfield, Greg G./Kinder, Kelly/Cherry, Kristin L./Daulton, Amy R., 2003: The West Wing as Endorsement of the U.S. Presidency: Expanding the Bounds of Priming in Political Communication, in: Journal of Communication 53, 427-443.
- Holtz-Bacha, Christina (Hg.), 2005: Europawahl 2004. Die Massenmedien im Europawahlkampf. Wiesbaden.
- Holtz-Bacha, Christina, 2006: Medienpolitik für Europa. Wiesbaden.
- Hooghe, Marc, 2002: Watching Television and Civic Engagement. Disentangling the Effects of Time, Programs, and Stations, in: Harvard International Journal of Press/Politics 7, 84-104.
- Imhof, Kurt, 2003: Politik im „neuen“ Strukturwandel der Öffentlichkeit, in: Nassehi, Armin/Schroer, Markus (Hg.), Der Begriff des Politischen. Baden-Baden, 401-417.
- Iyengar, Shanto/McGrady, Jennifer A., 2006: Media Politics. A Citizens' Guide. New York.
- Japp, Klaus J./Kusche, Isabel, 2004: Die Kommunikation des politischen Systems: Zur Differenz von Herstellung und Darstellung im politischen System, in: Zeitschrift für Soziologie 33, 511-531.
- Jarren, Otfried/Donges, Patrick, 2006: Politische Kommunikation in der Mediengesellschaft. Wiesbaden.
- Jarren, Otfried/Sarcinelli, Ulrich/Saxer, Ulrich (Hg.), 2002: Politische Kommunikation in der demokratischen Gesellschaft. Opladen.
- Johnson, Thomas/Kaye, Barbara, 2003: A Boost or Bust for Democracy? How the Web Influenced Political Attitudes and Behaviors in the 1996 and 2000 Presidential Elections, in: Harvard International Journal of Press/Politics 8, 9-34.
- Kaid, Lynda Lee (Hg.), 2004: Handbook of Political Communication Research. Mahwah.
- Kamps, Klaus/Nieland, Jörg-Uwe (Hg.), 2006: Regieren und Kommunikation. Meinungsbildung, Entscheidungsfindung und gouvernementales Kommunikationsmanagement – Trends, Vergleiche, Perspektiven. Köln.

- Kendrick, Alice/Fullerton, Jami A., 2004: Advertising as Public Diplomacy: Attitude Change among International Audiences, in: *Journal of Advertising Research* 44, 297-311.
- Kepplinger, Hans M./Maurer, Markus, 2005: Abschied vom rationalen Wähler. Warum Wahlen im Fernsehen entschieden werden. Freiburg.
- Klein, Markus/Rosar, Ulrich, 2005: Physische Attraktivität und Wahlerfolg. Eine empirische Analyse am Beispiel der Wahlkreiskandidaten bei der Bundestagswahl 2002, in: *Politische Vierteljahresschrift* 46, 263-287.
- Kleinfeld, Ralf/Zimmer, Annette/Willems, Ulrich (Hg.), 2007: Lobbying. Strukturen, Akteure, Strategien. Wiesbaden.
- Kleinstaub, Hans J., 2002: Auslandsrundfunk in der Kommunikationspolitik. Zwischen nationaler Kommunikation und Dialog der Kulturen, in: Hepp, Andreas/Löffelholz, Martin (Hg.), *Grundlagentexte zur transkulturellen Kommunikation*. Konstanz, 354-372.
- Klimmt, Christoph/Fischer, Till/Kuhrcke, Tim, 2005: „Mittendrin statt nur dabei“: Kriegsdarstellungen in Computerspielen, in: Knieper, Thomas/Müller, Marion G. (Hg.), *War Visions. Bildkommunikation und Krieg*. Köln, 276-292.
- Kriesi, Hanspeter, 2004: Strategic Political Communication: Mobilizing Public Opinion in „Audience Democracies“, in: Esser, Frank/Pfetsch, Barbara (Hg.), *Comparing Political Communication. Theories, Cases, and Challenges*. Cambridge, 184-212.
- Knieper, Thomas/Müller, Marion G. (Hg.), 2004: Visuelle Wahlkampfkommunikation. Köln.
- Korte, Karl-Rudolf, 2007: Mehrheiten für Unpopuläres in Deutschland? Bedingungen für eine erfolgreiche Reformpolitik und Reformkommunikation. Paderborn.
- Kreyher, Volker H. (Hg.), 2004: *Handbuch Politisches Marketing. Impulse und Strategien für Politik, Wirtschaft und Gesellschaft*. Nomos.
- Krüger, Udo Michael/Zapf-Schramm, Thomas, 2006: Sparten, Sendungsformen und Inhalte im deutschen Fernsehangebot. Programmanalyse 2005 von ARD, ZDF, RTL, SAT.1 und ProSieben, in: *Media Perspektiven* o. Jg., 201-221.
- Laver, Michael/Benoit, Kenneth/Garry, John, 2003: Extracting Policy Positions from Political Texts Using Words as Data, in: *American Political Science Review* 97, 311-331.
- Lester, Paul Martin, 2005: *Visual Communication. Image with Messages*. Belmont.
- Lilleker, Darren G., 2006: *Key Concepts in Political Communication*. London.
- Lilleker, Darren G./Lees-Marshment, Jennifer (Hg.), 2005: *Political Marketing in Comparative Perspective*. Manchester.
- Löffelholz, Martin, 2004: Krisen- und Kriegskommunikation als Forschungsfeld, in: Löffelholz, Martin (Hg.), *Krieg als Medienereignis II. Krisenkommunikation im 21. Jahrhundert*. Wiesbaden, 13-55.
- Lord, Carnes, 2006: *Losing Hearts and Minds? Public Diplomacy and Strategic Influence in the Age of Terror*. Westport.
- Machill, Marcel/Beiler, Markus (Hg.), 2007: *Die Macht der Suchmaschinen – The Power of Search Engines*. Köln.
- Maier, Jürgen/Brettschneider, Frank/Maier, Michaela, 2003: Medienberichterstattung, Mediennutzung und die Bevölkerungseinstellungen zum Euro in Ost- und Westdeutschland, in: Brettschneider, Frank/Deth, Jan van/Roller, Edeltraud (Hg.), *Europäische Integration in der öffentlichen Meinung*. Opladen, 213-233.
- Maier, Klaus, 2006: Newsroom, Newsdesk, crossmediales Arbeiten. Neue Modelle der Redaktionsorganisation und ihre Auswirkung auf die journalistische Qualität, in: Weischenberg, Siegfried/Loosen, Wiebke/Beuthner, Michael (Hg.), *Medien-Qualitäten. Öffentliche Kommunikation zwischen ökonomischem Kalkül und Sozialverantwortung*. Konstanz, 203-222.
- Marcinkowski, Frank/Meier, Werner A./Trappel, Josef (Hg.), 2006: *Medien und Demokratie. Europäische Erfahrungen. Media and Democracy – Experiences from Europe*. Bern/Stuttgart.
- Marr, Mirko, 2005: Internetzugang und politische Infomiertheit. Zur digitalen Spaltung der Gesellschaft. Konstanz.
- Marschall, Stefan, 2005: Idee und Wirkung des Wahl-O-Mat, in: *Aus Politik und Zeitgeschichte* 51/52, 41-46.

- Matthes, Jörg, 2004: Die Schema-Theorie in der Medienwirkungsforschung: Ein unscharfer Blick in die „Black-Box“?, in: *Medien & Kommunikationswissenschaft* 52, 545-568.
- Matthes, Jörg/Kohring, Matthias, 2004: Die empirische Erfassung von Medien-Frames, in: *Medien & Kommunikationswissenschaft* 52, 56-75.
- Maurer, Markus, 2003: Politikverdrossenheit durch Medienberichte. Konstanz.
- Maurer, Markus/Reinemann, Carsten, 2003: Schröder gegen Stoiber. Nutzung, Wahrnehmung und Wirkung der TV-Duelle. Wiesbaden.
- Mazzoleni, Gianpietro, 1995: Towards a „Videocracy“, in: *European Journal of Communication* 10, 291-319.
- McCombs, Maxwell, 2004: *Setting the Agenda: The Mass Media and Public Opinion*. Cambridge.
- McNair, Brian, 2003: *An Introduction to Political Communication*. London.
- McQuail, Denis, 2006: On the Mediatization of War: A Review Article, in: *International Communication Gazette* 68, 107-118.
- Medienpädagogischer Forschungsverband Südwest (Hg.), 2006: JIM-Studie 2006. Jugend, Information, (Multi-)Media. Stuttgart.
- Melissen, Jan (Hg.), 2005: *The New Public Diplomacy: Soft Power in International Relations*. Basingstoke.
- Merz, Manuel/Rhein, Stefan/Vetter, Julia, 2006: Wahlkampf im Internet. Handbuch für die politische Online-Kampagne. Münster.
- Meyer, Thomas, 2001: *Mediokratie: die Kolonisierung der Politik durch das Mediensystem*. Frankfurt a. M.
- Miller, David (Hg.), 2004: *Tell me Lies: Propaganda and Media Distortion in the Attack on Iraq*. London.
- Möller, Erik, 2006: Die heimliche Medienrevolution. Wie Weblogs, Wikis und freie Software die Welt verändern. Hannover.
- Müller, Marion G., 2003: *Grundlagen der visuellen Kommunikation*. Konstanz.
- Negrine, Ralph M./Mancini, Paolo/Holtz-Bacha, Christina/Papathanassopoulos, Stylianos (Hg.), 2007: *The Professionalisation of Political Communication*. Bristol.
- Neidhardt, Friedhelm, 2006: Akteure und Interaktionen. Zur Soziologie des Terrorismus, in: *Kraushaar, Wolfgang* (Hg.), *Die RAF und der linke Terrorismus*. Band 1. Hamburg, 123-137.
- Norris, Pippa, 2004: Global Political Communication: Good Governance, Human Development, and Mass Communication, in: *Esser, Frank/Pfetsch, Barbara* (Hg.), *Comparing Political Communication. Theories, Cases, and Challenges*. Cambridge, 115-150.
- Ott, Raphaela, 2006: Weblogs als Medium politischer Kommunikation im Bundestagswahlkampf 2005, in: *Holtz-Bacha, Christina* (Hg.), *Die Massenmedien im Wahlkampf. Die Bundestagswahl 2005*. Wiesbaden, 213-233.
- Peter, Jochen/Lauf, Edmund/Semetko, Holli A., 2004: Television Coverage of the 1999 European Parliamentary Elections, in: *Political Communication* 21, 415-433.
- Peters, Bernhard/Wefßler, Hartmut/Sifft, Stefanie/Wimmel, Andreas/Brüggemann, Michael/Kleinen-von-Königsloh, Katharina, 2006: Die Transnationalisierung von Öffentlichkeiten am Beispiel der Europäischen Union, in: *Leibfried, Stephan/Zürn, Michael* (Hg.), *Staatlichkeit im Wandel*. Frankfurt a. M., 230-261.
- Pfetsch, Barbara, 2003: *Politische Kommunikationskultur. Politische Sprecher und Journalisten in der Bundesrepublik Deutschland und in den USA*. Wiesbaden.
- Plasser, Fritz/Plasser, Gunda, 2002: *Globalisierung der Wahlkämpfe. Praktiken der Campaign Professionals im internationalen Vergleich*. Wien.
- Pohr, Adrian, 2005: Indexing im Einsatz. Eine Inhaltsanalyse der Kommentare überregionaler Tageszeitungen in Deutschland zum Afghanistankrieg 2001, in: *Medien & Kommunikationswissenschaft* 53, 261-276.
- Pontzen, Daniel, 2006: Nur Bild, BamS und Glotze? Medialisierung der Politik aus Sicht der Akteure. Münster.
- Raupp, Juliana, 2007: *Politische Meinungsforschung. Die Verwendung von Umfragen in der politischen Kommunikation*. Konstanz.

- Robinson, Piers, 2002: *The CNN Effect: The Myth of News, Foreign Policy and Intervention*. Oxford.
- Rössler, Patrick, 2006: Zur Logik der Agenda-Setting-Forschung, in: *Wirth, Werner/Fahr, Andreas/Lauf, Edmund* (Hg.), *Forschungslogik und -design in der Kommunikationswissenschaft*. Band 2. Köln, 139-167.
- Rössler, Patrick/Krotz, Friedrich (Hg.), 2005: *Mythen der Mediengesellschaft – The Media Society and its Myths*. Konstanz.
- Röttger, Ulrike (Hg.), 2006: *PR-Kampagnen. Über die Inszenierung von Öffentlichkeit*. Wiesbaden.
- Sarcinelli, Ulrich, 2005: *Politische Kommunikation in Deutschland. Zur Politikvermittlung im demokratischen System*. Wiesbaden.
- Saxer, Ulrich, 2007: *Politik als Unterhaltung. Zum Wandel politischer Öffentlichkeit in der Mediengesellschaft*. Konstanz.
- Scheufele, Bertram, 2003: *Frames – Framing – Framing-Effekte. Theoretische und methodische Grundlegung des Framing-Ansatzes sowie empirische Befunde zur Nachrichtenproduktion*. Wiesbaden.
- Scheufele, Bertram/Schünemann, Julia/Brosius, Hans-Bernd, 2005: Duell oder Berichterstattung? Die Wirkung der Rezeption des ersten TV-Duells und der Rezeption der Nachberichterstattung im Bundestagswahlkampf 2002, in: *Publizistik* 51, 399-422.
- Scheufele, Dietram A./Tewksbury, David, 2007: Framing, Agenda-setting, and Priming: The Evolution of Three Media Effects Models, in: *Journal of Communication* 57, 9-20.
- Schmitt-Beck, Rüdiger, 2004: Political Communication Effects: The Impact of Mass Media and Personal Conversations on Voting, in: *Esser, Frank/Pfetsch, Barbara* (Hg.), *Comparing Political Communication. Theories, Cases, and Challenges*. Cambridge, 293-322.
- Schmitt-Beck, Rüdiger/Schwarz, Frank/Abbaszadeh, Cyrus/Winter, Stephan, 2005: Wahlkommunikation im Internet. Eine Exploration zur Nutzung von „Informationslotsen“ am Beispiel des „WählerInformationSystems“ zur Europawahl 2004, in: *Zeitschrift für Parlamentsfragen* 36, 839-853.
- Schicha, Christian, 2003: *Die Theatralität der politischen Kommunikation. Medieninszenierungen am Beispiel des Bundestagswahlkampfes 2002*. Münster.
- Schoen, Harald, 2006: Der demokratische Musterbürger als Normalfall? Kognitionspsychologische Einblicke in die black box politischer Meinungsbildung, in: *Politische Vierteljahresschrift* 47, 89-101.
- Schoen, Harald/Weins, Cornelia, 2005: Der sozialpsychologische Ansatz zur Erklärung von Wahlverhalten, in: *Falter, Jürgen/Schoen, Harald* (Hg.), *Handbuch Wahlforschung*. Wiesbaden, 187-242.
- Schultz, Tanjev, 2006: *Geschwätz oder Diskurs? Die Rationalität politischer Talkshows im Fernsehen*. Köln.
- Schulz, Winfried, 1997: Political Communication Scholarship in Germany, in: *Political Communication* 14, 113-146.
- Schulz, Winfried, 2007: *Politische Kommunikation. Theoretische Ansätze und Ergebnisse empirischer Forschung zur Rolle der Massenmedien in der Politik*. Wiesbaden.
- Schweitzer, Eva J., 2004: *Deliberative Polling. Ein demoskopischer Ausweg aus der Krise der politischen Kommunikation?* Wiesbaden.
- Shah, Dhavan/Schmierbach, Michael/Hawkins, Joshua/Espino, Rodolfo/Donavan, Janet, 2002: Nonrecursive Models of Internet Use and Community Engagement Questioning whether Time Spent Online Erodes Social Capital, in: *Journalism & Mass Communication Quarterly* 79, 964-987.
- Shah, Dhavan/Cho, Jaeho/Eveland, William/Kwak, No jin, 2005: Information and Expression in a Digital Age – Modeling Internet Effects on Civic Participation, in: *Communication Research* 32 (5), 531-565.
- Snow, Nancy/Taylor, Philip M., 2006: The Revival of the Propaganda State: US Propaganda at Home and Abroad since 9/11, in: *International Communication Gazette* 68, 389-407.
- Strohmeier, Gerd, 2004: *Politik und Massenmedien. Eine Einführung*. Baden-Baden.
- Szukala, Andrea, 2005: Informationsoperation und die Fusion militärischer und medialer Instrumente in den USA. Der Versuch einer militärischen Antwort auf die neuen Bedrohungen in: *Medien & Kommunikationswissenschaft* 53, 222-240.

- Tenscher, Jens*, 2003: Professionalisierung der Politikvermittlung? Politikvermittlungsexperten im Spannungsfeld von Politik und Massenmedien. Wiesbaden.
- Tenscher, Jens* (Hg.), 2005: Wahl-Kampf um Europa. Analysen aus Anlass der Wahlen zum Europäischen Parlament 2004. Wiesbaden.
- Tiele, Annekaryn/Scherer, Helmut*, 2004: Die Agenda – ein Konstrukt des Rezipienten? Die Bedeutung kognitiver Informationsverarbeitung im Agenda-Setting-Prozess, in: *Publizistik* 49, 439-453.
- Tolbert, Caroline/McNeal, Ramona*, 2003: Unravelling the Effects of the Internet on Political Participation? In: *Political Research Quarterly* 56, 175-185.
- Voss, Kathrin*, 2007: Öffentlichkeitsarbeit von Nichtregierungsorganisationen. Wiesbaden.
- Vowe, Gerhard*, 2006: Mediatisierung der Politik? Ein theoretischer Ansatz auf dem Prüfstand, in: *Publizistik* 51, 437-455.
- Vowe, Gerhard/Dohle, Marco*, 2006: Politische Kommunikation, in: *Publizistik*, Sonderheft „50 Jahre Publizistik“, 250-275.
- Wagner, Jochen W.*, 2005: Deutsche Wahlwerbekampagnen made in USA? Amerikanisierung oder Modernisierung bundesrepublikanischer Wahlkampagnen 2005. Wiesbaden.
- Weaver, David H./Beam, Randal A./Brownlee, Bonnie J./Voakes, Paul S./Wilhoit, Cleveland G.*, 2006: The American Journalist in the 21st Century. U. S. News People at the Dawn of the New Millennium. Mahwah.
- Weischenberg, Siegfried/Malik, Maja/Scholl, Armin*, 2006: Die Souffleure der Mediengesellschaft. Report über die Journalisten in Deutschland. Konstanz.
- Weiß, Hans-Jürgen/Koch, Ansgar*, 2005: Wie Kriegsereignisse zu Medienereignissen werden. Die Ausstrahlung von Bildern gefangener und getöteter US-Soldaten im deutschen Fernsehen am 23. März 2003, in: *Knieper, Thomas/Müller, Marion G.* (Hg.), *War Visions. Bildkommunikation und Krieg*. Köln, 293-320.
- Weiß, Ralph* (Hg.), 2005: Zur Kritik der Medienkritik. Wie Zeitungen das Fernsehen beobachten. Berlin.
- Wimmer, Jeffrey*, 2003: Wie erfolgreich ist die politische PR von NGO's? Medienberichterstattung über den Weltwirtschaftsgipfel, in: *Donsbach, Wolfgang/Jandura, Olaf* (Hg.), *Chancen und Gefahren der Mediendemokratie*. Konstanz, 145-160.
- Wirth, Werner/Schramm, Holger*, 2005: Media and Emotions, in: *Communication Research Trends* 24, 3-39.
- Wirth, Werner/Schramm, Holger/Gehrau, Volker* (Hg.), 2006: Unterhaltung durch Medien. Theorie und Messung. Köln.
- Wolling, Jens*, 2006: Zur Logik der Mediamalaise-Forschung, in: *Wirth, Werner/Fahr, Andreas/Lauf, Edmund* (Hg.), *Forschungslogik und -design in der Kommunikationswissenschaft*. Band 2. Köln, 243-263.
- Zöllner, Oliver*, 2006: A Quest for Dialogue in International Broadcasting. Germany's Public Diplomacy Targeting Arab Audiences, in: *Global Media and Communication* 2, 160-182.