The Power of International Theory

Reforging the link to foreign policy-making through scientific enquiry

Fred Chernoff

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

The Power of International Theory

The field of international relations arose from the desire to assist and guide policy-makers to create a better and more peaceful world. However, many of the current trends, post-positivism, constructivism, reflectivism and postmodernism share a conception of international theory that undercuts the possibility of offering significant guidance to policy-makers.

The Power of International Theory critically examines these approaches and offers a novel causal-conventional alternative that allows the reforging of a link between international relations theory and policy-making. While recognising the criticisms of earlier forms of positivism and behaviouralism, the book defends holistic testing of empirical principles, methodological pluralism, criteria for choosing the best theory, a notion of 'causality' and a limited form of prediction, all of which are needed to guide policy-makers.

This book will be an invaluable text for advanced students and researchers in the fields of international relations theory and the philosophy of social science.

Fred Chernoff is Professor of Political Science and Director of the Program in International Relations at Colgate University, Hamilton, New York. He holds a PhD in Philosophy from The Johns Hopkins University and a PhD in Political Science from Yale University. He is also the author of *After Bipolarity*.

The New International Relations

Edited by Barry Buzan and Richard Little

London School of Economics and the University of Bristol

The field of international relations has changed dramatically in recent years. This new series will cover the major issues that have emerged and reflect the latest academic thinking in this particular dynamic area.

International Law, Rights and Politics
Developments in Eastern Europe and the CIS
Rein Mullerson

The Logic of Internationalism
Coercion and accommodation
Kjell Goldmann

Russia and the Idea of Europe
A study in identity and international relations
Iver B. Neumann

The Future of International Relations
Masters in the making?
Edited by Iver B. Neumann and Ole Wæver

Constructing the World Polity
Essays on international institutionalization
John Gerard Ruggie

The Continuing Story of a Death Foretold
Realism in international relations and international political economy
Stefano Guzzini

International Relations, Political Theory and the Problem of Order
Beyond international relations theory?
N.J. Rengger

War, Peace and World Orders in European History
Edited by Anja V. Hartmann and Beatrice Heuser

European Integration and National Identity
The challenge of the Nordic states
Edited by Lene Hansen and Ole Wæver

Shadow Globalization, Ethnic Conflicts and New Wars
A political economy of intra-state war
Dietrich Jung

Contemporary Security Analysis and Copenhagen Peace Research
Edited by Stefano Guzzini and Dietrich Jung

Observing International Relations
Niklas Luhmann and world politics
Edited by Mathias Albert and Lena Hilkermeier

Does China Matter? A Reassessment
Essays in memory of Gerald Segal
Edited by Barry Buzan and Rosemary Foot

European Approaches to International Relations Theory
A house with many mansions
Jörg Friedrichs

The Post-Cold War International System
Strategies, institutions and reflexivity
Ewan Harrison

States of Political Discourse
Words, regimes, seditions
Costas M. Constantinou

The Politics of Regional Discourse
Meddling with the Mediterranean
Michelle Pace

The Power of International Theory
Reforging the link to foreign policy-making through scientific enquiry
Fred Chernoff

Africa and the North
Between globalization and marginalization
Edited by Ulf Engel and Gorm Rye Olsen

Communitarian International Relations
The epistemic foundations of international relations
Emanuel Adler

Human Rights and World Trade
Hunger in International Society
Ana Gonzalez-Pelaez

The Power of International Theory

Reforging the link to foreign policy-making through scientific enquiry

Fred Chernoff

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2005
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Ave., New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group

This edition published in the Taylor & Francis e-Library, 2005.

“To purchase your own copy of this or any of Taylor & Francis or Routledge’s collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk.”

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Chernoff, Fred
The power of international theory: reforging the link to foreign policy-making through scientific enquiry / Fred Chernoff.
p.cm.
Includes bibliographical references and index.
1. International relations. 2. Policy sciences. I. Title
JZ1305.C44 2005
327.1'01-dc22 2004015955

ISBN 0-203-79946-1 Master e-book ISBN

ISBN 0–415–70138–4 (Print Edition)

Contents

Series Editor preface

The surge of interest in the study of international relations that occurred after the First World War was premised on the assumption that it must be possible, through systematic and scientific analysis, to promote a peaceful world. As we move into the twenty-first century, despite the violence and destruction that occurred in the twentieth century, for many scholars in the field of international relations, this conviction remains undimmed. Indeed, the experience of Europe over the past fifty years helps to sustain the belief that there is nothing utopian, in principle, about the desire to build a world that is both peaceful and prosperous. Most specialists in international relations accept, however, that there are still innumerable obstacles that make it impossible, at this juncture, to achieve universal peace and prosperity. As a consequence, policy-makers in countries across the globe are constantly being asked to make difficult international decisions that will have profound ramifications for the peace and prosperity of their own and other countries. This book asks if international relations theorists can, in principle, help policy-makers confronted by difficult international conditions, to establish sensible goals and to pursue policies that will enhance the chances of these goals being achieved.

What concerns Chernoff is the growing acceptance in international relations of meta-theoretical assumptions that create an unbridgeable gulf between international-relations theory and foreign-policy practice. The most damaging assumption, for Chernoff, is the one that rules out any possibility of prediction for international theory. If it is not possible for international theorists to say anything meaningful about the future, then it follows that they are unable to offer any assistance to policy-makers who wish to influence future events. Chernoff finds this aversion to prediction deeply puzzling, particularly in the case of critical theorists who are so committed to bringing about change in international relations. But Chernoff reveals that there is a growing number of theorists, fiercely committed to developing the study of international relations rigorously, who argue, nevertheless, that it is necessary to eschew prediction. Chernoff acknowledges that there is nothing inherently anomalous about this position, because it is accepted by some philosophers of science that there are areas of natural science where prediction is not possible. He accepts, as a consequence, that there may be significant areas of theory-building in international relations where prediction is also redundant. But at the heart of this project are the assertions that without predictions, theories hold little value for policy-makers and that there are no meta-theoretical reasons that preclude the possibility of prediction in the field of international relations.

To make this case, however, it is necessary to reassess key debates in the philosophy of natural science and to re-examine the relationship between natural science and social science. Chernoff favours an epistemological position that supports the view that theory choice in natural science and social science is determined, ultimately, by convention. Many objections have been raised in the philosophy of science to this position and Chernoff works through them, systematically, to show why they are flawed or can be overcome. The great strength of this epistemological position is that it can underpin both natural science and social science. At the same time, it reveals that social science has the potential for prediction, it promotes an open and pluralistic approach to theory-building, and it can also make room for normative international thought.

Chernoff concludes, however, that there are severe limits to what international-relations theorists can be expected to predict, and he insists that critics of prediction in international relations set the hurdle too high. Predictions, he accepts, can only be probabilistic, cannot extend very far into the future, and cannot involve many causal links. These are severe limitations. But even these relatively soft predictions can only be made if theorists in international relations engage in systematic research that contains a predictive dimension. Nevertheless, the point of principle remains. The fact of the matter is that most theorists in international relations are interested in the future and Chernoff provides them with firm philosophical foundations on which to stand and some clear guidelines for assessing the potential predictive value of their theories.

Richard Little

Preface

This book seeks to answer some foundational questions about the nature of the study of international relations. The field of IR was created after the First World War aiming to help guide policy-makers to bring about a better world, and such policy-guidance requires theory capable of some sort of grounding of prediction. The book focuses especially on trying to answer the question 'how might social science theories justify the sort of prediction needed for policy-making?' The question has seemed a vital one because many recent accounts appear to undercut the possibility of such expectations, which thereby undercut the original goals of the field of IR of helping to bring peace to the world.

In looking at these foundational questions I combed the published literature in international relations meta-theory and the philosophy of the social sciences hoping to find an account that would offer such a justification. None seemed to satisfy the standard criteria of adequacy, which led me to seek an alternative.

Eventually I also began to grasp the potential value of a fundamental insight that I encountered long ago as a graduate student in courses on the philosophy of mathematics and the philosophy of science, namely, that theory choice in physics (which includes a particular system of geometry) is inescapably and irrevocably conventional. From that point on, the two main tasks of this study became those of showing conventionalism's applicability to IR and the social sciences and finding solutions to various well-known philosophical problems with the most well-known varieties of conventionalism. The book thus argues that many of the problems of IR meta-theory that have received attention in the past fifteen years can be solved by applying specifically the form of the conventionalism that Pierre Duhem advanced in his works *The Aim and Structure of Physical Theory* and *To Save the Phenomena*. Whether the application of those principles to the social sciences is successful, as I contend, is of course for the reader to judge.

For permission to reprint passages from plays introducing the chapters I am grateful to the copyright holders. For the use of the passage by Euripides, I thank Ivan R. Dee; for Steve Martin I thank Grove Press; for Joe Orton, Grove Press in the US and Methuen in the UK; for Arthur Miller, Penguin; for Brian Friel, Catholic University of America Press in the US, and The Agency in the UK; for George Bernard Shaw, The Society of Authors; for Henryk Ibsen, The Gale Group; and for David Rabe, I thank Samuel French.

Over the (surprisingly large number of) years in which I developed the position taken here I presented parts of the whole at ISA meetings and on several college campuses. Without the many stimulating challenges from friends, colleagues, and students the arguments presented below would have been much weaker. For their comments and criticisms I would like to thank Al Yee of Georgia State, John Vasquez, David Stuligross, and Doug Macdonald of Colgate, Bruce Russett of Yale, and Ewan Harrison and Rod Hall of Oxford. Despite sometimes abbreviated exchanges, many of the most probing and stimulating challenges have come from Dan Nexon of Georgetown and Patrick Jackson of American University.

Grants and leaves from Colgate University have allowed me to finish the project much less slowly than I would otherwise have. Much help was provided by the library of the Yale Club of New York City, as well as Hunter College, New York University and the CUNY Graduate Center. I am thankful to several excellent student research assistants at Colgate, especially George Georgiev, Claire Putzeys, Aaron Sheldon, and Emily Weedon. Finally, for their enduring emotional support and encouragement without which this book would not be I would like to thank my friends in New York, David Frank, Steve DiFilippo, Marc Van De Mieroop, and Henry Yegerman, and my homies in Los Angeles, Dick Heller, John Aguilar, Kevin Merrill, Dusty Vinson, and Lee Arnold, as well as my sisters Myrna and Nastassja, my late brother Joel, the paterfamilias, Romo and, above all, my wonderful wife Vida, and her little entourage: Pato, Lydia and Monty the Dog. To Vida this book is dedicated.

New York, 1 June 2004

Abbreviations

CC: causal conventionalism
CS: conventionality of all science
d-n: deductive-nomological
DP: democratic peace
h-d: hypothetico-deductive
HT: hermeneutic tradition
IBE: inference to the best explanation
IP: incommensurability of paradigm
IR: international relations
i-s: inductive-statistical
RU: radical underdetermination of theory by evidence
SR: scientific realism

Introduction

SOCRATES: '… it is well-aimed conjecture which statesmen employ in upholding their countries' welfare. Their position is in relation to knowledge no different from that of the prophets and tellers of oracles, who under divine inspiration utter many truths, but have no knowledge of what they are saying.'

Plato, *Meno*, 99c, trans. W.K.C. Guthrie

The purpose of this book is to discover what value international relations (IR) theory holds for the making of foreign policy. Policy-making requires beliefs about what results a decision taken in the present is likely to produce in the future, and such beliefs require theory. One of the chief goals of the book, then, is to determine what, if any, sort of rationally justifiable beliefs about possible future events, that is, predictions, are justifiable in IR theory. The book offers a unique account of IR theory which will be termed 'causal conventionalism' (CC). In so doing, the book takes advantage of Pierre Duhem's revolutionary ideas about physical theory, first published exactly 100 years ago, which fundamentally altered the understanding of theory in the physical sciences, but which have been overlooked in the social sciences. This account has the ability to show that IR theory is much more powerful in its application than many current scholars seem to acknowledge.

Since the late 1980s, IR theorists have shown much more interest in these foundational and meta-theoretical questions, constituting what has been called the third debate. While the ultimate aim of this book is to provide something of value for the study of IR, which may then aid the practice of foreign policy, to do so requires examining questions in the philosophy of science and the theory of knowledge.

Answering the central question requires discovering what sort of structure IR theory must have to provide guidance to policy-makers. An understanding of the causal-conventional account allows one to improve the way in which theories of IR are developed, so that at least a part of the field will be relevant to and useful for policy-makers. Many studies of theory and policy in IR ignore the epistemic norms and proceed to offer theories or policy prescriptions. Similarly, many studies of meta-theory, particularly in the past twenty years, take one of various approaches (as will be shown) that exclude prediction – a key element needed for

IR theory to have relevance for policy-making. So the present book offers the causal-conventional account to show that IR theory can meet the epistemic norms of justifiable empirical enquiry and also have policy relevance.

The central questions – of the proper or appropriate character of IR theory and its value – are questions of meta-theory that can be answered only by considerations in the philosophy of the social sciences. There has been a debate in the study of the sciences, particularly since the publication of Thomas Kuhn's *Structure of Scientific Revolutions* (1962), as to whether our primary understanding of the nature and capabilities of science should be prescriptive or descriptive. That is, should the character and capacities of the sciences be studied by examining issues of philosophy and the norms of reason or by examining the history and actual features of scientific theories past and present? This book takes the view that a discipline – its scope and limits – can only be understood by an enquiry where history and philosophy are considered together; both must be examined in relation to one another. The inquiries into the philosophy of science and the history of IR theory reinforce each other in supporting the conclusion that, while certain epistemic norms are necessary, the field should strive for a genuine methodological and theoretical pluralism.

As this book seeks to answer the central question, it will have to determine what sort of claims IR theory makes. Do the claims constitute knowledge? If so, what sort of knowledge, and what may be done with it? If not, what is it that can be derived from the study of IR? Is it just inspired belief, as Plato suggests in the epigraph at the start of this chapter? The book will examine a number of foundational debates. Some of the debates have gained a great deal of momentum, such as the attacks on positivism by the range of post-positivist views, and especially the use of scientific and critical realism in IR (by Dessler 1989, Patomäki 2002, Patomäki and Wight 2000, Wendt 1999). Other issues raised here have been largely overlooked, e.g., the conventional component of scientific knowledge. (It should be noted that the attempt to offer epistemological 'foundations' for IR does not commit one to 'epistemic foundationalism': see pp. 57–8 and Chernoff 2002).

The book begins by considering what is involved in formulating policy. Chapter 1 focuses on the need to acquire rational expectations of future events and theories and it outlines the prominent theoretical and meta-theoretical positivist, post- and anti-positivist alternatives. The practical business of foreign policy-making demands beliefs, assumptions and conclusions about how a proposed policy will affect the world, which is a form of prediction. Prediction (at least of some sort) thus becomes a key to foreign policy-making. However, the notion of rationally grounded expectations of the future, or prediction, in the social sciences is quite controversial. Chapter 1 compares the rationalist accounts of IR to recent versions of constructivism and reflectivism. Rationalists regard much international behaviour as arising from given aspects of either human nature, the state or the international system. These features are treated as given outside of the explanatory theory of international politics. Constructivism, in contrast, argues that much more results from human decisions, choices and volitions, which might have turned out differently and which can change over time in ways that rationalists would deny. Constructivists treat agents and structures as distinct in concept but as inseparable in fact.

Chapter 2 begins the attempt to answer the question of what constitutes 'scientific knowledge'. Naturalism is the doctrine that the social sciences should be modelled, as far as possible, on the natural sciences – a view that will be examined throughout the remainder of the book. Chapter 2 asks the question as to just what constitutes 'natural science', a question one must answer in order to evaluate the claim of the naturalists.

Chapter 3 turns to the nature of the laws that theories of IR propose. Most philosophers of science distinguish the things we observe, which we refer to with 'observation terms', from the things that we do not directly observe but nevertheless appear to refer to in our theories by means of 'theoretical terms'. Justified inference about objects of the latter sort, their existence and their properties arises because of their role in theories that account for what we observe.

Chapter 4 focuses on causality in the social world, since it is generally regarded as the justification for the predictions policy-makers require. Causal notions may provide the rationale for connections between present conditions and future. These connections are not in general deterministic; they operate with somewhat weaker force. Accounts of the notions of 'probability' and 'probabilistic belief' are thus necessary for a full explanation of how the predictions involved in formulating policy may be justified.

Chapter 5 examines the arguments that critics of the scientific approach have levelled against 'prediction'. It considers three major lines of thought, each with a distinct source, objecting to the 'predictiveness' of social theory. It identifies flaws in each and offers a modified account of how prediction, broadly understood, is possible in IR.

Chapter 6 applies the causal-conventional account of meta-theory defended in Chapters 2–5 to two principal puzzles outside of the central cluster of questions for which the theory was composed, namely, (a) why is it that IR theory has, at least for the most part, failed to exhibit a natural-science-style approach-to-consensus (i.e., increasing agreement on new theoretical proposals – either by accepting or rejecting them) and (b) why has there been increasing agreement among scholars who debate the democratic peace (DP) hypotheses? On the first puzzle, the theory is shown to offer an account that fits with the pattern found in the discipline, and it offers the possibility, if not likelihood, of approach-to-consensus in the long-run. On the second puzzle, it is shown that what appears to be scientific-style 'progress' in the past two decades of debate over DP hypotheses is genuine and, furthermore, that it can be explained in terms of conventionalism.

The book concludes, in Chapter 7, with a summary of the power and capabilities, as well as the limitations, of the sort of knowledge that can be produced by the study of IR. The aim is to show how a body of rationally justified propositions, especially in the form of prediction, may be of value to the policy-maker. The book argues that there are important limitations on what can be known about IR and what can be predicted about the future; thus the most optimistic foundations for the study of IR are mistaken. But there is still a great deal that can be known. There are very meaningful and justifiable sorts of predictions. And policy-making, within certain limits, can be rational and effective. Thus the scepticism and relativism of many postmodern versions of reflectivism and constructivism are also quite mistaken.

Meta-theory is only one part of the study of IR and should not be taken to be more than that. Nevertheless, the importance of meta-theory cannot be denied. Elman and Elman (2002) offer the most recent plea for the pivotal role that meta-theory plays in the field of IR. They cite the surge of titles in IR literature that reference works by Lakatos and others in the philosophy of science, particularly in the area of theory appraisal. They point out that one's choice of research methods can have very practical consequences, such as winning research grants, gaining academic employment and publishing. So 'all IR theorists have an interest in the standards of appraisal' (Elman and Elman 2002: 232). They correctly observe (citing Bradley 1999: 316) that theory appraisal requires 'making explicit selections from among the menu of competing epistemologies. To refuse to engage in, and benefit from, methodological debate is to abandon the terrain to intuition and to the prejudices of whoever has the authority to decide the standards that should be applied' (Elman and Elman 2002: 233).

Because the aim of this book is to improve policy-relevance of the field of IR theory, the primary intended audience is students of IR theory, who must decide whether or not to develop predictive theories. It is hoped that the discussion below of the existing positions and the presentation of causal conventionalism will contribute to the foundational debate in some way, even if only by highlighting some of the alternative perspectives that have not heretofore received a fair hearing. But the book seeks also to introduce the reader to the questions and problems in the philosophy of the social sciences that students of international politics must confront, particularly in the current IR debates.

No particular background is assumed in the book beyond a basic familiarity with theories of IR. The book attempts to explain philosophical concepts sufficiently to clarify them for those unfamiliar with the details of the history of philosophy and the philosophy of science. It would benefit the reader to be acquainted with several of the excellent works on IR meta-theory, first and foremost Hollis and Smith (1991), Hidemi Suganami (1996) and the recent work by Heikki Patomäki (2002). While the exposition that follows is designed to provide all the necessary exposition for the student of political science who has no previous philosophical training, the primary goal remains to develop a causal-conventional account that answers the key questions of meta-theory and that offers a foundation for IR theory and foreign policy-making.

1 Policy-making, prediction and the theory of international behaviour

ACHILLES: 'What is a prophet? If he's lucky he gets one right out of ten. When his luck ends, so does he.'

Euripides, *Iphigenia at Aulis*, 1315, tr. Nicholas Rudall

The central question

National leaders have to make decisions that affect much of humanity. In the past few years President George W. Bush sent 130,000 American soldiers to Iraq, President Jiang Zemin has guided China into the World Trade Organization and President Vladimir Putin intensified the Russian offensive in Chechnya. Will these decisions lead to the desired outcomes? Can scholarly enquiry help leaders to make decisions that will produce the outcomes they seek? The latter question is the central focus of this book.

To find an answer, the book asks, further, what must IR theory be like if it is to have any value for policy-making, and if IR theory indeed has such a character? Policy-making requires, in one way or another, calculations or rationally justifiable beliefs about the future. Because the calculations must be rationally justifiable, they require the use of theory. An answer to the central question demands answers to questions about what inferences in IR are justified, especially inferences about the future, i.e., predictions. This is crucial for the policy-maker, who must judge what results are likely to follow from the choice of a particular policy. The policy-maker cannot offer judgements about what will result from which actions without a general set of principles that offer such links, i.e., without the adoption of a theory. So the book deals also with the question of how one chooses a theory and whether the best theories can be expected to be predictive.

Policy-makers and IR theorists ask questions about the causes and consequences of the US-led invasion of Iraq. What are the merits of invasion? Why was invasion the course chosen by the US administration? What are the prospects for a stable, democratic government in the wake of the invasion? Many recent works address these questions, such as Barton and Crocker (2003), Brooks (2002), Byman (2003), Feldman (2003), Heller (2003), Hollis (2003), Lewis (2002), Mearsheimer and Walt (2003), Metz (2003–4), Pollack (2002, 2003) and others. Some of these authors argue that if the US takes specified steps there is a reasonable prospect for a transition to a stable government in Iraq (e.g., Barton and Crocker 2003, Dobbins et al. 2003). Others are somewhat pessimistic about the chances for a transition to a stable government in Iraq (like Hollis 2003, Metz 2003–4). All of these authors offer generalisations, causal claims and predictions (discussed in Chapters 3, 4 and 5, respectively).

Most of these, and other authors who discuss theory and policy, tend not to discuss meta-theoretical and philosophical questions about formulating predictions in IR. There is thus a tension between this group and scholars who do discuss meta-theory, since the latter typically argue that prediction is not possible. The latter group includes IR theorists (Doran 1999, Gilpin 1981, Wendt 1999) and philosophers of social science (Bohman 1993, Little 1991, Taylor 1985, Winch 1958). This book advances a unique view, referred to as 'causal conventionalism' (CC), to resolve the tension.

It is not the direct aim of this book to answer policy questions such as whether the invasion of Iraq will reduce international terrorism. The aim is rather to show the proper procedures by means of which policy-makers may choose among policy options, including the means of justifying beliefs about connections between events. While questions of meta-theory are circumnavigated by many IR theorists, they are faced squarely by some contemporary authors (discussed in Chapters 3–6, see also citations by Elman and Elman 2002: 256–62) as well as by the most influential figures in the historical development of all major traditions in IR theory from Thucydides to Kenneth Waltz to Alexander Wendt.

Policy-making requires rationally based expectations about the future: predictions

Because policy-makers must make decisions, this book emphasises the need for predictive theory to formulate policy decisions. This need can be seen in the following way:

1 Policy-makers must make decisions.
2 Policy decisions require expectations about the future – a certain sort of justified belief about future events, which, broadly defined, constitutes, 'predictions'.
3 Predictions or expectations of this sort require beliefs about patterns of behaviour, that is, law-like generalisations.
4 Law-like generalisations are derived from or justified by theories (which typically have a causal element). Therefore,
5 Prediction-generating theories (among other things) are necessary for rational policy-making.[1]

Step 1 is a premise. This section defends step 2 and the next two sections defend steps 3 and 4, respectively. Step 5 follows deductively from 1–4. This chapter then will argue that the fundamental contribution that the academic study of IR can provide for policy-makers is to produce theories capable of yielding some sort of rationally grounded expectations about the future (i.e., predictions). The rest of this book probes the specific character and capabilities of those theories, especially their relationship to theories of natural science (Chapter 2), the nature of IR laws and observations (Chapter 3), causation in IR (Chapter 4) and the specific character of their predictive component (Chapter 5).

Two important qualifications are helpful here: one is that the argument presented below does not insist that IR theory is always useful to policy-makers, but only that *if* it is to be of use, it must include a predictive element (in the sense defined on p. 8). Second, this book does not argue that decision-makers always make rational decisions or draw upon IR theory. The argument claims only that *if* they seek to make rational decisions, then they must use IR theory. Leaders often make decisions without drawing on well-thought-out theories. But if IR theory is to aid them, then it must provide a rational basis for expectations about the future.

One can make simple decisions without a detailed theory. One 'predicts' that there will be, as scheduled, an election on the first Tuesday of November 2004. Although a prediction is needed, only a simple and uncontested set of causal generalisations are needed to produce that expectation. Nothing that needs to be honoured with the title of 'theory' is required. However, most foreign-policy decisions require more information and inferences than this example and policy-makers, in choosing an option to pursue, often must take positions on contested sets of nomic generalisations. More complex theories are needed in order to make those complex decisions rationally.

It is of course possible to make complex decisions without any good rational basis. No doubt that happens on occasion. But sometimes policies are indeed based on a fairly complex theory. For example, in the case of the Bush administration's decision to launch a war against Iraq, it does appear that its architect, Deputy Secretary Wolfowitz, in fact had a complex set of quite detailed causal beliefs about the relationship of democratic freedoms, democratic institutions, the steps necessary to create democracy in Iraq and the peace-inducing effects of an Iraqi democracy (Mufson and Ricks 2001). The theoretical analysis of Bernard Lewis is reported to have heavily influenced Secretary Cheney, Deputy Secretary Wolfowitz and their advisors. Presidential speechwriter David Frum is reported to have seen President Bush with a marked-up copy of a paper authored by Lewis (Waldman 2004).

President Bush and Prime Minister Blair believed that Afghanistan hosted al-Qaeda terrorist-training facilities and that Iraqi President Saddam Hussein possessed chemical and biological weapons which might fall into terrorist hands. Hussein's cooperation with terrorists rendered his weapons potential threats to the US, UK and other Western states. Bush believed several conditional propositions about the future, such as, (P1) *If al-Qaeda and its support-system are not physically destroyed, then al-Qaeda will likely continue to launch attacks on American citizens and American interests.*

Proposition P1 concerns the future; it says what is likely to happen at a time later than the moment of utterance. Second, it is implicitly based on a set of connected beliefs that are rational and arise from past observations. The beliefs may include 'terrorist organisations that are based on extremist ideology, that have advocated and committed violence, and that have access to weapons, training facilities and organisation opportunities, will continue to strike at those who they regard as their enemies unless stopped internally or externally'. Third, while the expectation is based on evidence that is regarded as adequate, the evidence need not be certain. Fourth, P1 is a conditional proposition, in that it says that the threat will materialise *if* certain conditions obtain. Fifth, the proposition does not guarantee that al-Qaeda will strike again, but only that such attacks are possible or likely. Hence, predictions may be either deterministic or, more typically in the social sciences, probabilistic.

With these features in mind, the term 'prediction' may be defined as follows: a rationally based expectation of the future, or prediction, in the natural or social sciences is a singular or general proposition which:

i) is indexed to the future relative to the moment of its utterance
ii) is based on a rationally justifiable body of theory, broadly construed
iii) may be based on imperfect evidence
iv) may be either deterministic or probabilistic and
v) may be conditional (i.e., of the form, 'if conditions C obtain, then result E will follow').

Even though most current meta-theorists in IR argue against social science prediction, many IR theorists have attempted to offer predictive theories. Some of those theories' predictions have been successful, one of the most ambitious of which is Bueno de Mesquita's *Predicting Politics* (2002). While many IR predictions have turned out to be correct, many have not. Thus the present book asks, is there a sound philosophical basis for believing that at least some sorts of predictions are rationally justifiable? Or, have correct predictions proven so only by luck (as Euripides says of the utterances of prophets, in this chapter's epigraph)? Few authors in IR have dealt with this crucial question. Among the exceptions are Patrick James (2002) and Ray and Russett (1996).

Some social scientists are very uneasy with the use of the term 'prediction'. However, if the term is used carefully, along the lines of the explicit definition given above, the usual difficulties should be obviated.[2] The negative reaction stems in part from the current rejection of positivist philosophy of social science, which endorsed a strong notion of 'prediction'. Positivism dominated much of the second half of the twentieth century and prediction was an important part of various positivist accounts of science. Positivist philosophers of natural science (e.g., Carnap 1956, Hempel 1962, Popper 1968, Schlick 1985) and social science (e.g., Hempel 1965, Neurath 1939, Popper 1945) emphasise the importance of prediction, especially in the corroboration (or as some preferred, falsification) of theories. As philosophers of science (e.g., Feyerabend 1978, Hanson 1958, Kuhn

1962 and Lakatos 1970) and the various social scientists attacked and discarded elements of positivism, they also – perhaps too hastily – threw out much of what positivists had to say about prediction.

The reaction against the notion of 'prediction' has gone so far that there are theorists and philosophers of social science who deny that prediction is possible in IR. For example, Charles Taylor, one of the most influential philosophers of social science, says that prediction in the social sciences is not only impossible, but is 'radically impossible' (Taylor 1985). Taylor, James Bohman (1993), Daniel Little (1991) and other philosophers offer a variety of reasons why social science theory is incapable of prediction. Criticisms of social science prediction are explored in detail in Chapter 5. Karl Popper also offers a well-known argument against a form of prediction in the social sciences (Popper 1957, 1971a, 1971b). Popper argues against historicism, that is, historical determinism and prophecy and its apparent alternative, 'utopian engineering'.[3]

Various IR theorists have also argued against prediction. For example, Donald Puchala contends that IR theory 'does not, because it cannot in the absence of laws … invite us to deduce, and it does not permit us to predict' (Puchala 1991: 79). Interpretivist and reflectivist IR theorists like Ashley (1986), Onuf (1989), Walker (1993) and others, following the lead of critical theorists and prediction-sceptic philosophers of social science, argue that IR theory (discussed in Chapter 3) is able to facilitate an interpretive understanding of events and deny that IR theory is capable of prediction or scientific-style explanation.

Even though many of these authors hope that IR theory can lead to 'human emancipation', their meta-theory undercuts its ability to do so. This trend in the theoretical literature in IR severs the link between IR theory and any significant ability to aid policy-makers to bring about emancipation or any other foreign policy goal. If they do not leave room for rationally grounded expectations about the future, that is, scientific-style prediction, then it will be impossible to formulate policies that can be expected to achieve various aims, including the emancipation of oppressed groups. Without the ability to say that a given action option has a higher probability than any of the other options of achieving the objective, e.g., a greater degree of emancipation of the target group, these theorists cannot recommend courses of action to achieve their desired goals. The loss of this essential capability has been largely overlooked by constructivists and reflectivists in the IR literature. All policy decisions are attempts to influence or bring about some future state of affairs. Policy-making requires some beliefs about the future, whether they are called 'expectations', 'predictions', 'forecasts' or 'prognostications'. The next step in the argument is to show how such beliefs can be justified.

Nomic generalisations as the foundations for rational expectations or predictions

How does one justify expectations, predictions or prognostications? Is there a difference between President Bush's believing P1, an astrologer's believing that he or she should avoid important financial decisions on his or her birthday and wishful beliefs, like the expectant parents believing that their child will make up for their own musical failings and grow up to be conductor of the Vienna Philharmonic Orchestra? In what way is President Bush's belief in P1 rational and a fitting basis for policy decisions, while the others are not?

Policy-makers, like most people, generally believe that the future will resemble the past. President Bush might accept the generalisation (G1) that *ideologically driven terrorist organisations continue to attack their avowed enemies until the organisations are stopped with external force or collapse from within*. They might also believe the generalisation (G2) that *ideologically driven terrorist organisations are most dangerous when they have (at least tacitly) the cooperation of states*. President Bush and his advisors believe, from observation, that these generalisations held true in the past and that they will remain true in the future. The generalisations, supported by past observations, seem to justify the prediction. But there is a serious philosophical question as to whether the generalisation may be believed to continue to hold in the future.

As Hume notes, people observe patterns in the past and seek to project them into the future. Hume argues that inductive reasoning typically justifies one's belief that the future will be like the past. He argues that inductive inference itself is unjustifiable. Many philosophers have sought to justify inductive reasoning in various ways. However some, like Popper, have concluded that induction is not philosophically justifiable and thus have sought to justify scientific inference without reliance on inductive reasoning. (This question is taken up in Chapter 4.)

While policy-relevant predictions may occasionally be justified by a single (often ad hoc) generalisation, this would be the exception rather than the norm. Most are justifiable only by recourse to a number of generalisations. For example, the prediction P1 that if al-Qaeda and its support-system are not physically destroyed, then al-Qaeda will likely continue to launch attacks on American citizens and American interests, may be justified by the generalisation G1 that ideologically driven terrorist organisations continue to attack their avowed enemies until the organisations are stopped with external force or collapse from within. But this does not justify a prediction about the benefits of a US-led invasion of Iraq. The prediction (P2) *American citizens and interests will be more secure if Saddam Hussein is deposed*, is more crucial to justify the policy of invasion of Iraq. The invasion, as a rational policy, may need several generalisations to support it. Examples would include, 'dictators do not voluntarily yield power (and thus continue to use all necessary means to ensure the survival of the regime)' and 'regional powers seeking hegemony come into conflict with global hegemons, if there are any'. These claims, combined with observations (e.g., Saddam Hussein is a dictator; Iraq seeks regional hegemony; the US is a global hegemon) justify the prediction P2.

At this point a difference emerges between President Bush's belief in P1 and the parents' belief about their unborn baby's musical destiny (but not the astrologer's belief). If the parents seek to formulate the generalisation with the most appropriate comparison groups (or reference classes), e.g., 'Children usually grow up to satisfy their parents' fondest hopes for them' or something of the sort. (This generalisation may look unfair but it is quite difficult to formulate another that would justify belief in the prediction and would appear probable.) Once it is stated, the parents would probably see that the generalisation is false. However, the astrological belief might not be distinguishable from Bush's belief in P1 on this basis, since some astrologers have extensive generalisations based on correlations of celestial observations with patterns of human affairs from which they derive their predictions. These generalisations are alleged to have empirical support and are the grounds on which they base their predictions. What then differentiates such beliefs from P1 as a basis for policy?

The link between generalisations and theories

Generalisations and causal mechanisms

Beliefs about the future, if they are rationally justified, must be based on something. The prediction that two brown-eyed human parents will (probably) produce brown-eyed progeny or that reducing trade barriers will (probably) raise living standards, if they are to be regarded as scientifically or rationally based expectations, derive from generalisations that are a part of a complex of generalisations and a systematic body of knowledge. It is not necessary at this juncture to produce a final list of features of a scientific body of knowledge. There are competing accounts that require careful study, which are considered in Chapter 2. What is important here is to recognise that there is a difference between the geneticist's expectation that the offspring of two parents will be brown-eyed and the parents' belief that the offspring will grow up to lead the Vienna Philharmonic.

Expectations or predictions (like P1 or the geneticist's belief), if rationally justifiable, are grounded in nomic generalisations. What are the grounds for the nomic generalisations? Three answers (considered below in Chapters 4 and 5) often cited are: that generalisations are tested by their functioning in a system of statements able to produce implications and explanations that each generalisation individually cannot; by their coherence with a complex of generalisations, singular statements and theoretical laws; and by their use of causal mechanisms.

As the example of prediction P2 shows, a complex of generalisations is capable of producing results greater than each generalisation taken separately. A researcher cannot hold only a single causal belief and hold in doubt every other experiential and theoretical proposition. Since other beliefs or knowledge-claims will be part of the accepted corpus, the generalisation in question must cohere with them. Internal consistency of beliefs is part of both empiricist and philosophical realist accounts. Duhem (1954), as is discussed below, holds that genuine testing is always testing of comprehensive bodies of theory-plus-background assumptions and is never testing of individual hypotheses or even of theories in isolation from the background and auxiliary beliefs.

Belief in the similarity of past and future patterns, i.e., the regularity of (physical or social) nature, is justified in part by the notion of 'causality'. Whether and how causality enters into social science reasoning is addressed in detail in Chapter 4. For present purposes it is important to note that causality, again, very broadly understood, provides an intellectual bond that holds together observed patterns and rationally based expectations about the future, such as the claim G1, G2 or (G3) that *politically unaccountable dictators who have behaved brutally toward their neighbours and adversaries over a period of decades tend to continue to behave brutally until they are forcibly deposed*. One's belief that something has the power or capacity to produce another thing is what justifies a belief that a specific future event (the ouster of Saddam Hussein) will, at least with some probability, result from a state of affairs the policy-maker sets up (large-scale invasion of Iraq).

Theorists invoke causal mechanisms to explain why these regularities occur. Consider the proposition uttered in 1991 (P3), that *Saddam Hussein will (probably) continue to repress the Iraqi people and attack neighbouring states*. Proposition P3 may be supported by different causal mechanisms. One sort of causal mechanism that gives rise to belief in P3 is a particular combination of the international and the domestic political systems. On this view, anarchy and self-help are the basic facts of international life, i.e., ordering principles of the international system. These features combine with many different types of domestic political systems over time and space, some democratic and some completely lacking internal political controls over excesses in the leader's aggression and brutality. When the states with the latter sort of domestic system are placed within an international system of anarchy and self-help, the result is that they exhibit extremely violent behaviour. On this (typically realist) view, states seek power and expand their power to the extent that they are able. Because in these cases there are no internal checks on tyrannical leaders, the leaders continue to act aggressively until prevented forcibly from the outside.

A second example of a causal mechanism that is sometimes posited pertains to the psychopathology of certain leaders. The US, UK and others employ social scientists to produce personality profiles of the various leaders (see Stephenson 1998). On this view, the policies of some autocratic states result from the attitudes and personal psychological make-up of the individual leaders. This is held to be especially true in the cases of states in which one individual is able to make major decisions with little interference from institutions or other government officials. The personality traits of these autocrats lead them to behave in certain ways. The behaviour exhibited in the past will continue more or less the same in the future, as long as the personalities are unchanged. The power to produce the subsequent events is 'causal power'. Thus it is important to consider (as Chapter 4 does) what sort of causal production, if any, occurs in international politics. In any case, if one is justified in believing, even probabilistically and conditionally, that some set of future conditions may come about, it is typically justified by recourse to causal powers. P1 may or may not be right, but it is the type of proposition that may legitimately be employed as a basis for policy, at least at this point, before criticisms have been examined.

The clusters of nomic generalisations and causal mechanisms, which are needed to generate predictions, constitute much of the substance of social and natural scientific theories. Predictions are needed if policy-makers are to have any rational basis for believing that particular policies will (likely) lead to desired outcomes. Thus predictions, and ultimately IR theories, are necessary for rational policy-making. At this point one can see the difference between President Bush believing P1 and the astrologer's believing in business setbacks on his or her birthday, since only the former is based on a cluster of generalisations, laws and postulated causal mechanisms. That basis provides the legitimacy for using P1 as a foundation for national policy and shows why the astrologer's prediction is not qualified to serve that purpose.

Resolving policy differences by theoretical appraisal

Perhaps the clearest way to see that expectations and the policies based on them must be grounded in theories is to consider disagreements. As in almost every country in every age, American presidential advisors are often at odds over the most effective means to solve problems. This is true of President Bush's cabinet, evident in the most important foreign-policy case, that of Iraq, as well as in the cases of North Korea and Afghanistan. On Iraq, Secretary of State Powell had a greater preference than other cabinet members for continuing diplomacy with allies and for taking military action, if needed, in a multilateral framework (Woodward 2002). Most security policy advisors in the administration accepted prediction P1 and the generalisation G1 on which it was based. Since substantial differences do arise among advisors, it is worth considering a hypothetical contrast of views.

In late September 2001 someone (though no one in the Bush cabinet specifically) might conceivably have predicted that (P4) *because most terrorist organisations seek public attention and the creation of fear in support of specific policy objectives, and because al-Qaeda has made such a monumental impact on the American public and has gotten the attention of the US administration and the world, al-Qaeda would henceforth be much more likely to demand that the US withdraw troops from Saudi Arabia, pressure Israel to alter policies and perhaps to aid Palestinians than to pursue further large-scale terrorist acts*. One who holds expectations along the lines of P4 might advocate policies other than invasions of Afghanistan and Iraq. A policy advisor who accepts P4 might recommend, rather, that the US begin to alter its Middle East policy, to withdraw military forces, to push for Israeli concessions to the Palestinians, etc. The difference in recommended policies could be connected as well to different views of the causes of the September 11 terrorist acts. There is a debate in the US between 'they hate us' and 'they hate our policies' as a diagnosis for the attacks. Those who accepted the latter view could envision negotiations or unilateral actions, like troop withdrawals from Saudi Arabia or ultimatums to Israel, that would satisfy al-Qaeda, assuage its hostility toward the US and eliminate its incentive for further terrorism against the US. In contrast, those who believe the problem is that 'they hate us' would see such concessions as producing no useful gain. The Bush administration did not believe the prediction P4 that al-Qaeda would be satisfied merely with changes in US Middle East policy or a role in negotiations. It believed rather that al-Qaeda would continue to attack American power and American interests.

The key here is to see why a policy of invasion should be preferred over US troop withdrawals, pressure on Israel, etc. The Bush administration advocated forcibly crushing al-Qaeda as far as possible because (though they had other differences) members of the administration all accepted prediction P1 and not P4, which in turn was because they accepted generalisation G1 and rejected (G4) that *ideologically driven violence-glorifying groups commit terrorist acts in order to gain negotiating leverage*. Why did they accept one generalisation and not the other? The answer is that only one of them, G1, coheres with the complex of connected beliefs that they hold about how the world operates and they hold that only one, G1, is based on a causal mechanism that the Bush security-policy advisors found plausible. A rational debate over whether invasion or negotiation would be more effective should include debate over whether P1 or P4 is more plausible; that debate will require a comparative evaluation of the theories in which they are grounded. Policy-makers seek analogous cases to compare with the case at hand and there are many academic theories about how this process works (e.g., Jervis 1976, Steinbruner 1974). But whatever the process, it must include the belief that past cases tell them something about the relationship between possible policy-action and probable outcome. The justification of the belief that those past cases shed light on the present and future includes belief in the generalisations that encompass both past and future events and the causal mechanisms that tie together policy action and outcome.

Some theories include G1 and the appropriate causal mechanisms as well as a complex of supporting generalisations, while competing theories include G4 and its causal mechanisms, along with other generalisations. Thus rationally based policies will be justified by rational debate over the most appropriate IR theory. If leaders deny that theories are important, they are still making decisions based on generalisations, laws and posited causal mechanisms – but they are using generalisations that they are refusing to acknowledge. In refusing to acknowledge them they are shielding them from scrutiny in the light of empirical evidence and is refusing to subject them to standards of coherence.

It is possible that an area of policy-making has no well-developed body of IR theory associated with it. So policy-makers may occasionally lack precisely appropriate IR theory. Still, leaders will have to form cause-and-effect beliefs about what policies will lead to what outcomes, and from there do the best they can. Though much progress is still to be made, most questions of war and peace have been heavily studied and theories have been offered as answers to them. The rational procedure, given that some decision must be taken, is to make use of the best evidence one can find, since even weak evidence is better than none.

The rational procedure is to make use of what is known and, depending on time-urgency and available resources, to construct a rough theory or proto-theory to guide the decision. Moreover, if IR theorists identify areas where decision-makers have to make important decisions in the absence of good evidence to support contending conditional statements of expectations (i.e., predictions), then IR theorists should recognise such areas as standing in need of further development. The point this chapter seeks to make is that if IR theory is to aid policy-making, then it needs to provide some basis for statements about expectations of the future and show that some statements are less worthy of acceptance than others. In the areas that IR theorists do debate, if they admit that their theories are incapable of providing a basis for conditional expectations about the future, then they must admit that their theories have little empirical relevance for policy-makers.

In sum, policy-makers must make decisions. President Bush decided to invade Iraq based on the objective of seeking to decrease the chances of further attacks on Americans. The policy was based on a (conditional and probabilistic) expectation about the future, i.e., a prediction, namely, P1, that if al-Qaeda and its support-system are not physically destroyed, then al-Qaeda will likely continue to launch attacks on American citizens and American interests. That prediction was rational because Bush accepted generalisations like G1 that ideologically driven terrorist organisations continue to attack their avowed enemies until the organisations are stopped with external force or collapse from within. The decision to attack Iraq required P1, along with a number of other beliefs about the future, such as P2, that American interests will be more secure if Saddam Hussein is deposed, but also predictions, such as that the costs would be affordable to the US, and the high-probability prediction that the attack would succeed in deposing Saddam Hussein. These predictions are justified by other propositions, while P2 is justified by the belief in the generalisation G2, that ideologically driven terrorist organisations are most dangerous when they have (at least tacitly) the cooperation of states in order to carry out major terrorist operations. Hence the policy of invasion, if it is rationally grounded, requires a complex of nomic generalisations and causal mechanisms that constitute a theory. The individual generalisations are worthy of belief only if they have a place in a theory, which is a coherent framework of generalisations, laws and causal mechanisms. Further considerations are adduced, based on hypothetico-deductive method, after the discussion of interpretivism on pp. 22–4 below.

Theoretical disagreement yields policy divergence

It is perhaps more difficult to see the connection between theory and policy-making in international politics than in economics or other social sciences. Nevertheless, such disciplines equally require that any theory that may be useful to policy-makers must generate rationally based expectations about the future. If it does not generate such expectations or predictions, then it cannot offer guidance to policy-makers except on the moral dimension. The systemic theories that most IR scholars debate are very general and ordinarily do not by themselves provide predictions that are of use to policy-makers in real-world problems. They offer only quite general predictions (e.g., 'states that do not socialise into the system will perish'). Some theorists, however, do offer useful predictions (Ray and Russett 1996). For example, one of the most intensely debated questions in IR, as Chapter 6 shows, is that of the DP claims. The most widely accepted hypothesis (which still does have some doubters), is that democracies rarely, if ever, fight one another. The acceptance of a liberal theory of international relations that includes this hypothesis would make a material difference to policy-makers. Its acceptance would lead to very different predictions than would its rejection and thus to different choices in the case of US policy toward the Middle East.

Consider two cases. First, the authors noted on pp. 5–6 above who discuss the conflict in Iraq offer different policy prescriptions because they adopt different theories. A classical realist might argue that the use of force, invading Iraq and ousting Saddam Hussein, will advance US national interests because it would limit the threat that an otherwise-dangerous Iraq would be able to pose. Such theorists, like Pollack (2002), have argued that as Iraq acquired more weapons of mass destruction it posed a growing threat to Israel, to other neighbours, to regional stability and to US interests in the Middle East. Pollack has argued as a political realist that Saddam Hussein should be removed from power for security reasons – and that Iraq's violations of UN Security Council resolutions should be regarded as a lesser factor.

Other analysts who reject political realism have argued that an invasion decreases Western security. Rosemary Hollis (2003), for example, makes the case that any decrease in the threat to US security from terrorists that Saddam Hussein might have sponsored or contributed to will be overshadowed by the increased threat to US security from the violent anti-American sentiment that the US invasion of Iraq will stimulate. Hollis has argued that a US invasion of Iraq will send alarm-signals to states in the Middle East. It will ominously foreshadow how Washington will exercise its power in the Middle East in the coming years and thereby produce a backlash. 'Islamist groups will portray [US behaviour] as a clash of civilizations, and in the chaos that would ensue from the toppling of existing regimes anti-American sentiment is more likely to spread than to give way to secularist democracy and free markets' (Hollis 2003: 26).

Authors like Pollack (2002) argue that war is in the best interest of the US, while those who reject realism, like Hollis (2003), argue that it is not. There are two different and incompatible policy prescriptions (war/no war) that are aimed at maximising the goal of US and Western security. They differ in part because they are based on different law-like generalisations, which are parts of different theories. In order to achieve the desired policy goal it will be necessary to choose the most effective policy option, which may be accomplished only with the most adequate theory available. This book thus must seek to shed light on the factors that should properly guide the choice of a theory.

The second case is that in which the Bush administration desires peace between Israel and its Arab neighbours, the administration presumably believed that Saddam Hussein supported terrorism (e.g., his government offered payments to the families of Palestinian suicide bombers in Israel and the occupied territories). Given the DP theories (advanced by Doyle, 1983a, 1983b, Levy 1989, Owen 1997, Rummel 1983, Russett 1993 and others), a democratic Iraq is more likely to live in peace with Israel than a non-democratic Iraq. (The same holds true for a democratic Palestinian state.) So from the point of view of US interests, it is worth expending considerable resources to change the regime in Iraq from that of Saddam Hussein's dictatorship to one of democracy. Theorists who oppose DP hypotheses (like Gowa 1999, Layne 1994, Spiro 1994 and others) would not accept such a claim.

While various sorts of opponents of DP hypotheses might agree with the Bush administration that there are other good reasons to work to overthrow Saddam Hussein, they would not insist on creating democratic institutions in Iraq and might be satisfied with a regime type closer to that of other US allies, such as Egypt, Jordan, or those of various recent Latin American and Asian non-democratic allies. Of course, any policy option has to be considered in terms of the material, human, moral and legal costs. With respect to the type of Iraqi regime that would be most preferred by the US, policy-makers who support DP theories in IR would advocate democracy, while policy-makers who adopted the theories of DP critics would not regard democracy-building as especially or uniquely favourable to peace. These two groups advocate different theories, the theories generate different predictions, and the predictions incline policy-makers to adopt different courses of action.

Predictive, non-predictive and anti-predictive theories

Some reflectivist, interpretivist and postmodern, as well as some rationalist IR theories are anti-predictive in that they deny the possibility of rationally grounded expectations about the future. This study argues that these anti-predictive IR theories undercut the ability of the discipline to aid policy-making. There are, though, some IR theories that are non-predictive without being anti-predictive, and thus do not threaten the link to policy-making. There are at least two major categories of such theories, namely, highly general empirical theories and non-empirical, normative theories. With regard to the first group, as just noted, some IR theorists present principles or laws at a high enough level of generality or abstraction that they do not appear to be able to generate predictions that are precise enough to be of use to policy-makers. Two points are relevant here. First, Waltz (1986: 6) and various other IR theorists do endorse the importance of prediction in theoretical pursuits. While their theories do not generate specific, policy-useful predictions, they do generate predictions. Thus these authors are not prediction-sceptics. Empirical theories at a high level of generality are typically not anti-predictive. In many instances they are reasonably seen by their authors as an element in a scientific complex of description, explanation and prediction, which can aid policy-makers.

A second category of IR theory that is not in and of itself predictive, but is not anti-predictive, is moral and normative theory. IR theorists often apply moral arguments and precepts to problems in foreign policy and international relations. Those theories are essential as guides to action. No matter how much policy-makers may know about causal relationships among political, economic and social factors, they must have some set of objectives they seek in order to choose one option over the others. They must know if what they seek above all else, or to a particular degree, is national economic expansion, pursuit of global human or political rights, protection of citizens from external attack, expansion of national political influence or military power, etc. The objectives are generally supplied in part by theories with moral content. But the policy-maker must also have the best idea possible about what steps and strategies are most likely to lead to those goal states, which can be supplied only by theories with empirical content and predictive capabilities. Bueno de Mesquita and Lalman (1994) present one clear example of the use of general principles of IR, along with much more specific observations, to produce a predictive theory. (See also Bueno de Mesquita 2002.)

Many contemporary philosophers deny that a clear and definitive distinction can be drawn between the empirical and the moral. But the distinction drawn here between empirical theories and normative theories – or theories with empirical and normative *components* – does not require a philosophically impregnable distinction between the two. A rough distinction will suffice. One might deny a hard-and-fast 'fact–value' distinction but assert that the value-content of a statement like 'Slave owners are morally reprehensible' differs greatly from that of 'Patrick is taller than Lydia'. The naturalist meta-theory defended here opens up a role for theories that are descriptive, explanatory and predictive. But such theories do not close off a role for moral theories. Indeed, moral principles, imperatives and theories require both descriptive and predictive theories, since 'ought' implies 'can'. One must know what is possible and probable in order to make or appraise foreign policy decisions. One may not morally condemn the lifeguard for rescuing only one of the two drowning swimmers if it was physically impossible to save both. Similarly, in international politics one may not blame a state dedicated to just and egalitarian democratic rule for not creating a just and egalitarian order throughout the system, if that state does not have the resources to do so. One must have causal and descriptive theories in order to understand what the state has the capacity to do, even when it comes to appraising how well it lives up to its moral obligations. Policy-making unavoidably requires both theories that are primarily moral and theories that are primarily empirical. This book endorses a proper role for normative theory in world politics in appraising past decisions and in choosing the best future courses of action, though it focuses on the primarily empirical form of theory. Theories of IR inspired by critical theory and postmodernism are on stronger ground when they offer strictly normative arguments.

One might object that the emphasis here on the need for empirical theories and the need to know consequences would be vulnerable to charges that it endorses ethical consequentialism, which some philosophers reject. But the position here is clearly not that actions are to be evaluated as morally good or bad in terms of their consequences. Even anti-consequentialist positions generally recognise that a moral agent must have knowledge of conditions to perform moral deeds, even if those actions are appraised on the basis of the agent's motivations rather than the actions' consequences.

Finally, a comment is in order about the relationship between theories and predictions. Theories in the natural or social sciences do not themselves contain predictions. 'Predictive theory' is a term used here to refer to those theories that, when combined with factual or non-theoretical knowledge, are capable of generating predictions.[4] In most natural sciences and social sciences, theories themselves are not tied to any temporal moment. They include nomic generalisations that may be applied, if at all, at any particular moment in time. The generalisations refer to kinds of events or conditions, but the theories themselves do not have any intrinsic link to a given moment of history and do not provide dates when those conditions will obtain. The theory may or may not contain principles that apply to the conditions that obtain at a given moment. If they do, either in a positive or negative way, then the theory may be used to generate positive or negative (probabilistic and/or conditional) predictions. If the theory includes the principle 'unpopular dictators do not relinquish power without the use of force', it is not clear without an examination of conditions at the moment whether the leader of state X is a dictator, is unpopular, and is the target of the use of force. If the leader of state X is an unpopular dictator and is not the target of force, then the *application* of the theory, which makes use of an empirical study of current conditions that is not itself part of the theory, will help to generate a prediction (e.g., that that the dictator will not relinquish power).

Rationalist and constructivist approaches to IR

The rationally grounded decisions policy-makers take will be based on, among other things, a theory of IR. Theories vary in level of generality, from high levels of abstraction, like those of Kaplan (1957), Keohane (1984), Organski (1958) and Waltz (1979), to theories of particular regions, states or individuals. The welter of general theories of IR may be divided into categories along many different dimensions, such as those dealing with the most important variables (like higher versus lower level of analysis, material versus ideational variables); the nature of the actors (fixed or exogenous versus endogenous identities and interests); inside versus outside (interpretivist-hermeneutic versus scientific-style); the characteristic behaviour of actors (maximisers of relative gain versus absolute gain); the types of methods the theories use (qualitative versus quantitative); or the nature and aim of the theory (whether they are primarily explanatory/descriptive versus primarily prescriptive).

Opposing types of theories need not always be mutually exclusive. It was argued above that explanatory/descriptive and/or predictive theories should be combined with normative theories to produce policy decisions. The two are of different types and a researcher may adopt theories of both types. Similarly, there are elements common to inside and outside theories and one might adopt both inside and outside theories to answer different questions.

IR theorists ask a variety of different sorts of questions, such as:

1 Do democracies fight fewer wars than non-democracies?
2 Are bipolar systems more stable than multipolar systems?
3 Was the US nuclear alert in 1973 intended as a warning signal to Brezhnev?
4 What did Winston Churchill mean by an 'iron curtain'?
5 Is the use of nuclear weapons ever justified?
6 Are preventive wars immoral?
7 Should leaders ever follow policies that do not advance national interest?

The appropriate methods for each of these questions differ. Some of these questions (1–2) are clearly best answered by methods that would include an examination of many cases and an attempt to find associations among variables; some (3–4) are most effectively answered by means of interpretive methods; and others (5–7) are best answered by the analysis of concepts and the application of moral theory. There is, moreover, no single, unified IR or social science 'field equation' into which all must ultimately be made to fit, though there are some specific forms of reasoning that are applicable to all. Because the social world is complex and multifaceted, even questions about a possible war between the US and North Korea yields questions that, though related, sometimes require orthogonal and cross-cutting theoretical approaches and methods. Wendt notes that rationalist and constructivist theories need not be seen as entirely opposed; each seeks to answer a different set of questions, neither set of which is 'more important' than the other (Wendt 1999: 34).

Rationalists take the interests, preferences and identities of actors as given, while constructivists do not (Wendt 1999: 33–4). For rationalist theories of IR, the goal of self-preservation or maximising interests or power, for example, would qualify as pre-established elements of the identities of states; the rationalist theory of IR would then explain how the state rationally pursues those interests. Constructivist theorists argue that the identities of the actors change over time and that the changes are affected by aspects of the international political system. The general approaches to social theorising that constructivism and rationalism cover include a broad range of theories. Both political realist and liberal theories, as well as most Marxist theories, are generally understood as rationalist, while reflectivist, interpretivist, postmodern and poststructural are generally regarded as constructivist.

The various highly general theories of IR provide parameters within which mid-level theories would have to operate; the combination of theories at different levels and empirical study of conditions at a particular moment in time provides the basis for the predictions that give rise to policy decisions. How much a policy-maker should rely, for example, on cooperative reciprocity in shaping international organisations or on the use of force will depend on what sort of theory is accepted, whether its nature is realist, institutionalist, postmodern, Marxist or other.

Constructivists emphasise that classical realists and liberal institutionalists both see the state as a rational actor, with pre-established interests and preferences, and as seeking to maximise gains in an anarchic international system. Constructivists group both traditions together along with Marxism as 'rationalist' and offer an alternative, according to which utilities are assigned as a result of interactions; the states or agent-entities cannot be construed as existing apart from one another or apart from the system of interactions of which they are a part.

Some constructivists (Wendt 1999, Wight 1996) develop a scientific-style approach while others (Ashley 1986, Kratochwil 2001) adopt a more postmodern or interpretivist constructive account. While constructivism is a general approach to social theory, specific constructivist theories of IR, like Wendt's, are sometimes seen to offer a middle ground between political realism and liberal institutionalism (see also Jackson and Nexon 1998).

Constructivists see individual states and the international system as not entirely distinct. Nations or states develop interests and social identities, which change over time. But at any given moment, they inhabit an environment of shared understandings and expectations. Social reality is the product of people constructing that reality (Onuf 1989: 35–65). The identities of the states or ethnic, state or national groupings and the identities of the systems they inhabit constantly interact with, create and re-create one another.

According to constructivists, an 'identity group' is formed over an extended period of time through historical processes, where the identities are transmitted from generation to generation through a process of socialisation, though some evolutionary changes are introduced along the way. The people within that group self-identify as part of it. National groups, such as the English, Armenians, Japanese or Kurds, are formed in this way. States often – but not always – coincide with national identities, but other important aspects of identity, such as religion, race and class, will form groupings that exist across state boundaries. Sex groupings, of course, transcend every state border and constitute an important identity-grouping aspect (Tickner 2001). Objective factors such as these are important in developing the identity groupings. But they are not sufficient. Constructivists hold that the element of 'social construction' of these groupings is necessary for a full account of how they arise.

A crucial difference between constructivism and rationalism is that the latter, especially political realism, sees anarchy as a defining, constant and essential feature of the modern international system. Constructivism holds that there are no such essential and constant features: all are mutable. Anarchy exists because historically it evolved along a specific, contingent path in world politics. Another system with the same power-configurations among its members might have evolved with a more or less cooperative sort of relationship.[5]

Constructivists ask where the interpretations of the world come from and how they influence behaviour. They see institutions as coming from a process of cognition that shapes the social environment. The environment in turn shapes cognitive processes. Constructivists point out that norms produce changes in typical patterns of state behaviour. For instance, human rights norms have occasioned intervention in the past decade in a way unexpected by political realists, who emphasise a system that places a premium on sovereignty and respect for the internal autonomy of states.

Many constructivists point out that rationalists, especially political realists, overlook justice in IR, since political realism neither provides an account of 'justice' and 'injustice' in international regimes and institutions nor shows how the activity of theorising itself reflects the interests of the powerful groups in society. Constructivists point out that political realism's amoral emphasis on power, which avoids questioning the legitimacy of that power, acquiesces to the power relations within society, even though those power relations may be unjust.

In the midst of increasing globalisation of many aspects of international political and economic relations, there is also greater factionalisation and disintegration, which constructivists say can only be explained by a focus on the self-identification of groups, a focus that differs from political realists' focus on material factors. They see large-scale change as much more problematic than political realists or liberal institutionalists do. Constructivist theories admit that both material and ideational factors shape behaviour. For political realists, cooperation in international politics takes place, when it does, in a system of self-help, anarchy and mutual distrust, while for constructivists there is the possibility of altering the pure self-help nature of the current system and reducing the mutual distrust to create a more cooperation-prone system, even in the midst of anarchy.

Inside and outside theories

Constructivism is a form of what has come be known, following Collingwood (1946) as 'inside' theories. These contrast with 'outside' social science theories that make use of the approach of the natural sciences. The controversy in IR is sometimes known as 'The Third Debate', as it followed the realist–idealist debate in the first half of the twentieth century and the traditionalist–behaviouralist debate of the third quarter of the century. The outside approach had dominated during the post-war period, with many scholars in IR, political science, economics and sociology working within their disciplines on the assumption that there are clear law-like connections between different factors, which can be identified and quantified.

If one wants to develop a theory of the formation or expansion of alliances, one might study factors that are believed to be most likely to have an influence on alliance behaviour, such as similarity of ideological orientation, geographical proximity or distance, power balances, conventional arms races and nuclear arms races. One examines, compares and quantifies examples from history. On the basis of the statistical or comparative tests, one might conclude that power balances, geographical location and conventional arms races do contribute to the formation of alliances, while ideology and nuclear arms races do not. This outside method would constitute one step in building a theory of international alliances. The outside approach does not require that any questions be asked about the meaning or symbolic value of an alliance or the personal or psychological traits of the national leaders. It does not at any point require the investigator 'to get inside' the events by seeing them from the point of view of the agents or to learn the vocabulary or psychological framework of the decision-makers.

The outside, causal method is modelled on the style of enquiry typical in the physical sciences. An outside explanation in IR will invoke some sort of theory and law-like regularities, in ways considered below in Chapter 3. The outside perspective includes theories whose key explanatory variables are at various different levels of analysis: they may refer to individual human agents, to their roles in bureaucracies and governments, to political, economic or sociological characteristics of the individuals in large groups (states, societies), to the characteristics of the international state system(s), or to other levels.

The inside approach challenges the outsiders' claim that causal, scientific-style explanation is even possible in the social sciences. The inside approach sometimes uses the term 'human sciences' to emphasise that the goal is not to provide an account of something as abstract as 'society', but of complex human beings, who are full of meanings, motives, symbols and intentions. The inside perspective offers an understanding of events and ways to interpret them (or to show their constitutive elements) rather than to identify a causal chain of events.

How would a supporter of the inside approach offer advice to policy-makers, given that such an individual holds that social science theories are non-predictive? The insider argues that theory in IR, as with other social sciences, is capable of explaining observed events in a way that allows an understanding of them, but cannot generate scientific-style explanations or rationally grounded expectations of the future.

Suppose that after 11 September 2001 the US Secretary of State, who recommended an invasion of Iraq, had been an interpretivist who was taught that IR theory could at best decode the meanings of actions but could not offer scientific-style explanations and could offer no rational basis for expectations about the future, that is, predictions. Instead of working out policies based on predictive theories, the Secretary of State says the US will simply push forward pursuing specific goals, whatever they may be – US national security, protection of Western values, universal human rights, etc.

A defender of predictive theory could point out that the pursuit of any goal requires that one must have some beliefs about what actions will likely lead to the advancement of those goals. Any claim that a particular policy will have a particular effect is an expectation about the future, or a prediction, in the broad sense defined on p. 8.

If asked about the grounds for the decision, the Secretary of State would say that predictive IR or social science theory was in no way a guide to action. When asked whether the Secretary of State believed that an invasion with a force of 180,000 American and British troops would succeed in ousting the regime of

Saddam Hussein, he or she would say 'yes'. When asked why he or she believed this, the answer would probably be something like 'past invasions have been successful when they consisted of a technologically superior force moving against a degraded defensive force, whose leadership was viewed by its citizenry as unpopular or even illegitimate'. The 1991 Persian Gulf war may be cited to illustrate some aspects. The Secretary of State might next be asked why a past case would justify the belief that a 180,000-strong invasion force would very likely succeed. It would quickly become clear that the leader, if rational and not choosing the policy purely randomly, believes a set of conditional propositions of the form, 'if policy c is followed, results e will probably follow'. If probed as to why each of the conditionals is true, the Secretary of State would probably offer a causal explanation, invoking mechanisms that pertain to the destructive capabilities of certain weapons and well-trained troops. In short, the policy-maker would support his or her policy-decisions with causal generalisation and would be going beyond the limits of hermeneutic-interpretive recommendations. Without a causal theory, the invasion decision would not be justifiable.

Many generalisations are needed to justify a belief regarding the launching of an invasion (or creating a free-trade area, expanding a security alliance like NATO, etc.). Are the generalisations mutually contradictory? If so, rational policies will be difficult to formulate and execute. When the generalisations are considered separately, tested against their competitors, and combined to produce a consistent picture of world politics, the result is a theory of IR.[6] Leaders may make decisions based on a theory that is explicit, tested against alternatives, and scrutinised for internal inconsistencies, or they may make decisions while denying that any theory is at the foundation. In the latter case, they are either genuinely making the decisions randomly with no foundation, or they are implicitly making use of a set of connected general and causal beliefs, that is, a theory, or proto-theory. If, however, they use theory implicitly, they are doing so without subjecting them to scrutiny, in which case they are much more likely to include some generalisations that conflict with one another and others that could, if explicated, be proved false. Policy-making is irrational if choices are made randomly. It is less rational than it might be if it is based on theories that are kept implicit and permitted to escape study and scrutiny.

If the future is not predictable, then the above argument shows that theories of IR (as well as economics and other social sciences) cannot offer guidance to policy-makers outside of purely moral advising, as the discussion above of the erstwhile postmodern Secretary of State suggests. Chapter 5 takes up the crucial question of whether it is possible to make predictions or hold rational expectations about future social behaviour – and thus whether IR is really capable of aiding policy-making; it takes up also the question of whether the hermeneutic tradition is perhaps more capable of prediction than its proponents think, given that one form of inference crucial to hermeneutic reasoning is hypothetico-deductive (Hempel 1988: 1–3).

One important point should be added here regarding the relationship between the inside and outside approaches to the study of IR, which is that they should not be regarded as mutually exclusive. In much the same way Wendt characterises rationalism and constructivism (see p. 20), the opposition between the inside and outside approaches presented in Hollis and Smith's (1991) influential text has perpetuated the notion of their incompatibility (see Pollins forthcoming). The methodological pluralism advocated in this study not only finds room for both to co-exist, but emphasises the positive need for both sorts of theories. One of the bases of the methodological pluralism here is the recognition of the wide array of questions asked by scholars (p. 20 cited several). The question 'Are bipolar systems more stable than multipolar systems?' can best be answered by means of scientific-style theories, while the question 'Was the US nuclear alert in 1973 intended as a warning signal to Brezhnev?' can best be answered by means of an interpretive analysis. Each question in IR should be answered by the methods and theoretical structures best suited to it. Consequently, the diversity of questions in IR demands different sorts of theories, scientific, interpretive, normative, etc.

Criteria of theory choice

How a leader chooses the most effective response to challenges from state-supported terrorism, nuclear proliferation, a declining trade balance or a currency crisis will depend on how he or she assesses the character of the states involved, the proclivities of particular leaders of those states, as well as on the general theory of IR he or she adopts. But on what bases does the leader choose a theory?

This study will approach the question: how does the theorist and policy-maker choose between the competing theories of liberalism, political realism, etc.? It must be remembered that meta-theory alone does not entail a unique theory or allow the investigator to determine which theory is correct, though it offers help in the form of guidelines and criteria. Meta-theory should be able to help the scholar and policy-maker identify improperly formulated theories – theories that do not meet the criteria of adequacy – and help them to ascertain appropriate criteria of theory choice by means of which they may determine, in conjunction with a body of empirical evidence, which theory is to be preferred. As noted, a high-level theory like structural realism will not always be specified precisely enough to pick out a unique option from among a large number of policies under consideration. If there are only a few policies on the table (e.g., bomb or negotiate), then there is a higher chance that one of the various theories will lead to a unique choice. But often there are many choices under discussion, e.g.:

i) bomb military targets isolated from civilians
ii) bomb military targets located near sparse-to-medium density population areas
iii) bomb all military targets
iv) negotiate with a small set of options
v) negotiate with many options, and so on.

When there are few choices the theory will likely be inconsistent with some and it may not dictate a unique option among the rest.

As the book considers several IR theories, it will therefore defend various criteria for how one should go about choosing between the competing theories. The discussions of criteria in this book make use of two important distinctions: that between internal and external criteria and that between intrinsic and comparative criteria. Internal criteria have to do with the characteristics of a theory, and external have to do with the relationship between the theory and the external world, especially observations of it. Intrinsic criteria are characteristics of a theory, which might include both sorts of features just noted. Comparative criteria deal with the relationships between a theory and its rivals, i.e., how a theory stacks up against competitors. Most philosophers argue that theories are ultimately only to be compared against rival theories – to be replaced only when better theories emerge. Just because the old theory has broken down, investigators still do not have as one of their rational or scientific options eschewing theory altogether. Despite its sputtering and clanking, a previously accepted theory will continue to be used until a more smoothly operating new theory replaces it.

Traditionally, the most hallowed internal criterion is self-consistency, followed closely by the criteria of parsimony (i.e., economy or simplicity) and explanatory range (i.e., the ability to account for a wide range of phenomena). One might add another, the ability to account for phenomena in greater detail. With respect to parsimony, rational investigators prefer, other things being equal, a theory that has few laws, variables or theoretical entities over a theory that has many. Likewise, given two equally simple theories, rational investigators prefer the theory that explains more types of observations or allows explanation and prediction of more specific sorts of events.[7]

A problem with even a very small number of desiderata is that they sometimes come into conflict with one another. Suppose that there are just two theories, where one is simple but explains a limited range of phenomena, while the other explains far more types of phenomena, or explains them in greater detail, but requires more variables or laws to accomplish this task. Without trying to settle this difficult issue at this stage, it is helpful to consider the extreme cases. Consider three theories, one of middling complexity and explanatory range, a second that is extremely parsimonious but does not explain the world in detail at all and a third that has immense explanatory range but has many, many variables and laws. If the latter two theories cannot be applied by policy-makers, then there is a basis for choice in this case. Whatever the limitations of the first theory, it provides guidance for policy-makers. Since the policy-maker must make a decision, only the first theory provides a rational basis for that choice among policies.

One might object that there are many grounds on which to choose a theory, and policy-applicability is only one. This is entirely true. But to the extent that IR seeks to affect the real world and aid policy-making, then the latter two theories are not options for the policy-maker. Furthermore, if a critique of prediction-generating theories were mounted and sustained, then one would have to accept that IR theory cannot aid policy-making in any empirical way. Absent powerful philosophical grounds for rejecting the predictiveness of social science theory, then IR theory should be constructed so as to aid them, both on empirical and moral dimensions, given that policy-makers have to make decisions.

A final comparative criterion to note is that of allowing for the least possibility of error, that is, the least likelihood of producing false conclusions. This will require trade-offs against the criterion of maximising the range of phenomena that are accounted for; a theory that minimises the likelihood of false consequences may well account for a narrower range of observations than one that accounts for a larger range of observations but opens one up to the possibility of greater error. How these various trade-offs should be treated will be considered in the chapters that follow. (The specific question of criteria for theory choice is raised again in Chapter 3, see pp. 78–84.)

The conventionality of all science

Conventions and scientific theory

This book defends a version of conventionalism, a view generally overlooked in contemporary IR theory and methodology. It will be shown that the version of conventionalism usually identified by critics is not the only or best version. An alternative, modified form of conventionalism is able to provide the needed meta-theoretical guidance. (Perhaps a different appellation would help ease confusion, such as 'causal conventionalism' or the less euphonious 'quasi-Duhemianism'.) In order to allay possible distress on the part of the reader, some comments on the sort of conventionalism that is, and is not, endorsed here are in order.

There are several distinct varieties of conventionalism. The version defended here is quite specific, distinct from any previous version, and eschews many of the most commonly known aspects of conventionalism. The term 'conventionalism' is retained because the fundamental insight is that there is a conventional element to all systematic empirical knowledge. This study assumes that physical theory is the best example of a systematic and useful body of knowledge, that physical science is as secure and certain a body of knowledge as any empirical enquiry and that other fields, unless otherwise demonstrated, are not immune from limitations that affect physical theory.

Since there are various conventional constraints on the objectivity of the theories of the physical world, similar constraints (and perhaps others) affect the study of IR and the social sciences. But the conventionality of all science (CS) thesis enables the philosopher of the social sciences to explain many phenomena in the real world and in the history of the study of IR that are often explained only by invoking much more limiting accounts of knowledge, such as Kuhn's incommensurability of paradigm (IP) thesis or the radical underdetermination of theory by evidence (RU) thesis (discussed further in Chapter 6).[8]

The CS thesis defended here holds that there is a conventional element to all scientific knowledge as evidenced in physics and chemistry. The decision to adopt any scientific theory requires conventional choice between competing postulates (such as the measure-stipulation in physics; see Chapter 6). But the scientific process of making this choice is subject to dispute, a process that will ultimately produce clear grounds for choosing one postulate over the others. That is, there is no reason to assume that, because a conventional, extra-theoretic choice must occasionally be made, such a choice must be arbitrary or based on amorphous or ambiguous considerations such that it makes no difference, for the investigator's purposes, which convention is chosen. Several arguments will be adduced in Chapter 4 seeking to establish that conventionalists may offer precise, rational and unambiguous guidelines for choosing among conventions.

Conventionalism maintains that part of what many philosophers take to be objective or factual knowledge is a matter of agreement among speakers, practitioners or investigators in a discipline – though the principles that guide agreement are not arbitrary. Conventionalism also has particular relevance to the special issues surrounding the notion of 'necessity' or 'necessary truth'. But this point is often over-emphasised by critics (as described below).

What conventionalism does not entail

In examining the literature on conventionalism, especially in Chapters 5–7, two points should be kept in mind. First, if some form of conventionalism is refuted, it does not necessarily follow that all forms are false. Second, if some proponents of conventionalism are found to make use of faulty arguments against its rivals, it does not follow that the conventionalist is incapable of formulating successful arguments against the doctrine's rivals.

Simple linguistic conventionalism points out that it is merely a matter of agreement among speakers that the words employed to refer to things are chosen for that purpose, and other assignments might just as well have been made. The word 'penguin' might have been used to refer to pickles and vice versa. The statement 'penguins are vertebrates' is true only as a matter of convention. Had 'penguin' been chosen to refer to pickles, the statement would have been false. But this is a trivially true point about language and does not capture anything important about the epistemological doctrine of conventionalism. It is a mistake, as will be shown, to dismiss conventionalism on the grounds that it includes nothing more than trivial linguistic observations of this sort.

Conventionalism in the theory of knowledge is often conflated with scientific constructivism and it will be shown in Chapter 6 that the two are quite distinct. Kuhn has argued that scientific knowledge is constructed, is relative to a paradigm, and that there is no connection between paradigms. He thus supports the IP thesis. Ultimately, there is no way to adjudicate competing knowledge claims because there is no theory-neutral language that would allow an unbiased, or even an intelligible, comparison. The paradigms themselves are in 'a sense … constitutive of nature' (Kuhn 1962: 110). The differences between

IP and CS are described in more detail in Chapter 6. Suffice for the moment to note that conventionalism is different from the 'epistemological constructivism' of Kuhn, often referred to as the historical or sociological account of science, in that conventionalism holds that there are linguistic and conceptual bases for direct comparisons of competing theoretical frameworks and that there are rational grounds for choosing one theory over its rivals.

Conventionalism differs from scientific realism (SR) in that the former does not assign to all the claims of science either true or false status (see Chernoff 2002). The conventional element is present and should not be forgotten entirely by scientists. In his classic paper 'Empiricism, Semantics and Ontology', Carnap (1956) argues that pragmatic considerations are important in science and that they can help to guide the scientist in a choice of theory or a theoretical language. Conventionalism has strong negative implications for essentialist metaphysics, since it holds that metaphysics studies only what one's conventions are and how they may be mistaken for essential features of objects, species, etc.

Conventionalism is sometimes taken to identify 'analytic truth' both with '*a priori* truth' and with 'necessary truth'. This view had many adherents until the advent of semantics for modal logic, which allowed philosophers to interpret statements of necessity in a way that did not entail that they are analytic or *a priori*. Philosophers in the 1970s came to accept that there are necessary *a posteriori* truths. However, conventionalism is not limited to the simple identification of the necessary with the *a priori* and/or the analytic. This view does not provide an answer to the question of the source of modality. Alan Sidelle (1989) has argued for an anti-realist account of 'necessity' and notes that there have been few realist defences of the necessary *a posteriori*, perhaps because philosophers implicitly accept that the existence of necessary *a posteriori* propositions entails that one must accept a philosophical realist account of 'necessity'. Sidelle correctly argues that there are no grounds for any such supposition; a realist interpretation must be defended and not assumed.

A related criticism of conventionalism is that it cannot be reconciled with the rejection of theories in science. Critics argue that conventionalism maintains that the truths of a theory are essentially definitions and thus cannot be refuted by experience and observation. Hence, no observation will suffice to tarnish a theory and the advancement of observational techniques will never undermine the acceptability of a theory once it is adopted. According to these critics, the history of science, where theories have indeed been shown to be inadequate and overthrown in favour of new rivals, shows the untenability of this position. This criticism is addressed in Chapter 4, where the differences between different versions of conventionalism are delineated. The merits of Duhem's form of conventionalism, which does not treat scientific laws as analytic statements, are highlighted.

Social sciences and conventionalism

The version of conventionalism defended here is quite distinct from the traditional formulation found in the philosophy of the natural sciences. Many of the differences will be adumbrated in the chapters that follow. The philosophical merits of CS and Duhemianism in the philosophy of the natural sciences also translate to a considerable extent to the social sciences. But the view defended here shares elements with meta-theoretical approaches that are generally quite far from CS. For example, causal conventionalism (CC) shares with older forms of empiricism the goal of prediction; with critical realism, the views that causal explanation is a part of the social sciences and that reasons should be viewed as causes with critical realism; and with reflectivism an emphasis on the problem of subjectivity.

This view also embraces methodological pluralism. The social world is complex and multifaceted. There are many avenues and many directions from which to approach it. There are some limits, as various pseudo-academic approaches lack legitimacy and philosophical grounding. Still, many approaches are possible – quantitative, qualitative, social scientific, humanistic, interpretive, moral and legal and others. A view of the IR and social sciences consistent with Little's (2000) interpretation of the English school is strongly endorsed here.[9] This study, consequently, argues against any single approach, rational choice, statistical modelling, postmodern or constructivist, that rejects any role for all others. While a form of naturalism is defended in the chapters that follow, this only accounts for a subset of all issues and research questions in the social sciences; other issues and research questions require other sorts of methods. Especially to the extent that the social scientific enquiry can aid policy-making, it must be recognised that questions of ethical norms, moral values, equity and legal status will have to be confronted in the formulation of all policy. As noted on p. 18, even for those who deny a philosophical basis for the fact–value dichotomy, there are vast differences in the factual or evaluative content of different sorts of statements.

Conclusions

Those who argued in favour of a policy of nuclear deterrence had beliefs about the future, given the adoption of certain policies of weapons deployment. Those who argued that certain types of leaders or regimes cannot be deterred but will act aggressively in the face of certain types of nuclear deployments likewise had various beliefs about the future, given the adoption of particular policies of weapons deployment. No one in the deterrence debate could avoid making probabilistic claims about the future without being willing to argue that at least some future actions are predictable. This book argues that there are, indeed, non-logical but rationally defensible criteria for choosing a theory. These meta-theoretical elements will have a substantial impact on what connections one ultimately sees between present actions and future outcomes.

Policy-making, if it is to be a rational enterprise at all, cannot proceed without expectations of how present actions will affect future outcomes. The expectations that arise from a policy proposal will depend on what sort of theoretical orientation the policy-maker accepts. It might be tempting to eschew the explicit adoption of any IR theory. But, as shown above, if such a tack is taken, the policy-maker will still be basing policy on unstated beliefs about how a given action will affect future outcomes. Thus one would be, essentially, ‘smuggling in’ some sort of unstated theory. And by denying recognition of the reliance on theory, one forecloses the option of refining and improving the accepted beliefs about the efficacy of different sorts of options to bring about different sorts of results. So whatever theory the policy-makers adopt, they will, either implicitly or explicitly, lead to the connections seen between present actions and future outcomes. The policy-maker must then choose between the various competing theories.

The acceptance of rigorous criteria for theory choice should not, as just noted, be taken as precluding methodological pluralism or an approach that endorses multiple and overlapping theories. This study endorses both. In the case of methodological pluralism, it is crucial to use the methods appropriate for the study at hand and the data available to answer the questions as fully as possible. Some problems allow the use of hundreds or thousands of cases to be applied to find a solution or answer the central research question. In other cases there will be few, if any, (relevantly) comparable cases, so that statistical analysis will be useless. Likewise, short-term or agent-based questions may demand lower-level or less general theories for their analysis. Such theories may complement or overlap more general theories, which are more appropriate for the pursuit of different sorts of questions. This study seeks to answer some but not all questions of IR meta-theory. The book does not insist that CC is a solution to all problems of meta-theory; there are other fruitful ways of getting at various problems. The insights from the philosophy of the physical sciences, especially the CS thesis, help advance the understanding of IR meta-theory. But insights from other disciplines, such as the biological sciences and hermeneutics, are of great value and are considered below.

This chapter has argued that policy-making requires some form of prediction. Many traditional accounts of scientific method treat ‘prediction’ as unproblematic. Philosophers, especially since Newton, have raised various objections. But ‘prediction’ remains a part of standard accounts of the method of (most) natural sciences. Adjustments have been made, e.g., the scrapping of symmetry with explanation. However, unlike in the natural sciences, in the social sciences the objections continue to be raised.

This study focuses on the question of what value IR theory is to policy-makers. So the most important preliminary question to answer in the course of the study is what sort of prediction, if any, is justifiable in IR. By any measure, and certainly in comparison with other fields of study, the past four centuries of physical theory have been overwhelmingly fruitful and successful. Physical science has been used to change the world. An account of IR theory that mirrors the theoretical and predictive character of modern physical theory is of much greater value to the policy-maker than an account that eschews the physical-science model and predictiveness, even if it does provide a clear basis for some (e.g., hermeneutic) form of 'explanation'. If the philosophical objections to any form of predictive naturalism are insuperable, then they must be admitted. But it is a goal of this study to produce a (partially) naturalistic, predictive account, if such an account is philosophically justifiable.

2 Social science, naturalism and scientific realism

RANCE: I am a scientist. I state facts. I cannot be expected to provide explanations.

Joe Orton, *What the Butler Saw*, Act II

PARRIS: We are here, Your Honor, to discover what no one has ever seen.

Arthur Miller, *The Crucible*, Act III

Those who analyse foreign-policy options draw a connection between a policy option, like the invasion of Iraq, and future conditions, like a decreased threat to US security. How do they make those connections? Must they posit causal relationships of the sort familiar in the natural sciences between the policy options and the expected results? Can policy-makers treat foreign-policy decisions as deriving from theories that resemble natural-science theories? Those who believe that theories of IR and other social sciences are fundamentally similar to natural-science theories are regarded as 'naturalists'.

The analogy between natural science and social science

The natural sciences have transformed human affairs so profoundly that few can help being struck by their power, utility and sweep. This has been especially true of the physical sciences, like astronomy, chemistry and physics. Because many natural sciences have been so successful, their methods and principles have been applied elsewhere, with the idea that the legitimacy of the natural sciences will then, according to the doctrine of 'naturalism', be conferred upon the area to which the methods are applied, whether it is the study of suicide patterns or national security.

The impact of Newtonian physics was so great that thinkers in many fields came to believe that the application of similar, precise (mathematical) methods to their disciplines would produce similar results, namely, the creation of a single theory with straightforward laws capable of subsuming a great variety of observations. It would reduce divergence and dissent within the field and usher in a single view that would be accepted by consensus. So thinkers in many fields embraced the model of the physical sciences. In the seventeenth century, Spinoza developed his system of ethics using scientific systematisation; he drew on the geometrical method of Euclid, which had recently been incorporated by Newton into his physical theory. In the development of the social sciences in the seventeenth and eighteenth centuries, Leibniz believed that human behaviour could be brought into a Newtonian-like system. Scholars in the nineteenth century, such as Comte (1970), constructed a positivist sociological theory.

The approach adopted here takes seriously the idea of a scientific study of international politics. The central focus of this study on requirements of policy-making, which includes predictiveness, motivates this approach and tends to favour naturalism. A science of international politics should have at least some of the capabilities of the natural sciences. Naturalism is pushed to a recognised limit. This approach is not intended to exclude all possible alternative approaches (unless they explicitly contradict it). There is no attempt to deny that policy-making also requires considerations of equity and moral obligations. While there may be no complete and absolute fact–value dichotomy, the study of astronomy and the study of ethics nevertheless emphasise facts and values very differently. Thus authors who advocate the application of the methods of humanities disciplines and 'value theory' for IR are encouraged to do so.

One of the differences between the natural and the social sciences, according to non-naturalists, is that only in the former do the best theories have genuinely true laws, while in the latter, the social universe is so complex and has so many variables that come into play across a range of cases that even the best theories have laws that, if they can offer any useful guidance at all, are idealisations or approximations and are thus not literally true. The generalisations of Waltz regarding balancing of states or of Keohane regarding the cooperation-inducing flow of information within regimes, may have value but they are imprecise approximations or idealisations. This book defends a fairly strong version of naturalism and in Chapter 3 evidence is offered attempting to show that the idealisations of the social sciences are akin to and not a contrast to the laws of the natural sciences.

This chapter begins with a brief review of the context of the debate over scientific methods, which saw positivism emerge as dominant, to be later challenged by a range of criticisms. It then moves on to the central task, that of laying out cases for and against naturalism. The first half of this chapter lays out the positivist view of science through much of the twentieth century, which was liberally applied in the social sciences in the middle and latter part of the century. The second half of the chapter offers anti- and post-positivist critiques of this view, beginning with attacks on the main thrust of the programme, its account of the natural sciences (pp. 39–42). Next are the attacks on the positivist account of the social sciences, which are more central to the debate between naturalists and non-naturalists in IR. Those attacks are of two sorts, criticisms by post-positivist authors who share with positivists a scientific-causal approach to the social sciences (pp. 42–44) and anti-positivist attacks by authors who reject the 'outside' scientific understanding of the social sciences and instead endorse the 'inside' interpretivist or hermeneutic account of the social sciences (pp. 44–51). The chapter then turns to the critique proposed by critical or scientific realism (pp. 51–58) and some responses (pp. 58–60).

The tradition of positivist explanation

The positivist account of science

The earliest theories of the twentieth century were reactions against the dominant neo-Kantian and Hegelian idealism that mixed metaphysics with science. Early twentieth-century empiricists sought to dismiss unobservable entities as metaphysically unfit for scientific duty. Some even held that explanation was not a proper part of, but rather a kind of psychological adjunct to, the central work of scientific theorising (reflected in the quotation from Joe Orton in this chapter's epigraph).

Within a few years, logical empiricists and logical positivists, whose influence grew quickly, came to include explanation in their account of science. Through the second half of the twentieth century, IR adopted standard social-scientific conceptions of 'explanation' and 'prediction'. The dominant approach was quasi-scientific, though often also seeking causal or law-like regularities among types of events, which, along with accounts of mechanisms and bridge principles, constituted theories.

Logical positivism was developed in Vienna in the 1920s and 1930s. The view spread through the English-speaking world during and after the Second World War and came to dominate the Anglo-American scene through the 1960s. Positivism focused on the observable and objective and sought to distinguish authentic knowledge from all other sorts of purported knowledge, largely using the model of natural science for that purpose. Logical positivists sought to identify and eliminate that which lacks genuine meaning (and only appears to have meaning), including metaphysics and moral, social and religious philosophy. Early positivists emphasised verification and verifiability as the key to meaningfulness, later preferring 'confirmability'.

Logical positivism, which later evolved into logical empiricism, led by Reichenbach (1938), advocated the hypothetico-deductive (h-d) model, wherein the proffered hypothesis is subjected to empirical testing (see also Salmon 1979). Deductive consequences are derived from the hypothesis and compared to the results of the empirical tests. While there were hostile critics in the Frankfurt School, more sympathetic critics of logical positivism included Karl Popper and Carl Hempel.

Karl Popper criticised logical positivism and especially its verification criterion, preferring instead his criterion of falsifiability. He rejected positivism's search for 'meaningfulness' advocating instead a criterion for the demarcation of science from non-science. But Popper shared some of positivism's views and sought to account for both natural science and social science.[1] Popper attempted a demarcation criterion that would separate genuinely scientific discourse from discredited forms of enquiry, such as astrology and other fields that were still taken to be scientific, especially Freudian psychology and Marxist dialectical materialism.

Carl Hempel is usually classified with the positivists who came onto the scene after the first wave of positivism had crested. He worried about the positivists' treatment of 'confirmation'. Hempel elaborated the covering-law model in the form of the deductive-nomological (d-n) model, which many positivists came to endorse. According to the covering-law model, scientific explanations are deductive arguments in which the explanandum is a conclusion and the explanans is the set of premises. This d-n model treats explanation and prediction in similar ways; the two are the same except for the time-frame of the minor premises and conclusion. Hempel explicitly extended the covering-law model to the social sciences through his development of probabilistic variant of the d-n model, the inductive-statistical (i-s) model (Hempel 1962). For Hempel, explanations do not merely terminate the investigator's desire to pose questions, they provide an objective, empirically verifiable ground to answer the question.

Positivists supported the fact–value dichotomy, according to which factual and evaluative statements can be neatly distinguished from one another; science should proceed by dealing with the former and entirely eschewing the latter. And they vigorously supported the observation–theory dichotomy, which states that observation statements could be understood as based on theory-neutral experience, while statements involving theoretical terms could not (Carnap 1966: 225–31).

Naturalism: the unity of science

The traditional view of natural science is that its methods have succeeded in bringing investigators to a veridical picture of the natural world. It seemed to follow that those methods, applied in areas other than natural science, would have the best chance of bringing investigators in other disciplines to the truth. The central question of this chapter, pursued also by many logical positivists, is whether the study of world politics and of human social structures is like the study of nature.

While some of the most prominent members of the group, especially Moritz Schlick and Rudolf Carnap, were interested primarily in physical science and especially in the newly developed system of relativistic physics, others, like the sociologist Otto Neurath, pursued the implications of the positivists' notion of the 'unity of science' for the social sciences and the life sciences. Popper and, especially, Hempel were closely associated with the positivist movement and were primarily interested in accounting for natural science but also wrote significant treatises on the social sciences.

Logical positivists agreed that basic positivistic principles of knowledge applied wherever one seeks knowledge, and that the fundamental methods that had been so successful in the physical sciences should be applied elsewhere. The methods of all sciences were one and the same, and thus science was unified by its methodology. Hempel, Neurath and others who focused attention on developing theories of the social sciences advocated 'the unity of science', i.e., the view that essentially one and the same scientific method should be applied to the natural and social worlds. A similar notion is that science, whether physical, biological or social, has an inner character in virtue of which it is a science, and that there is a single method that allows the investigator to bring out that character. The view here that science has such an essential nature, and that all sciences share it, is sometimes referred to (usually by its critics) as 'scientific essentialism'. The term is often used as a synonym for 'naturalism'.

Eight characteristics of natural science

Naturalism argues that the social sciences share the basic character of science with natural sciences like physics, chemistry, biology, and so on. So it is helpful to summarise the dominant – largely positivist – account of the foundations of science, which Salmon calls 'the received view' (1989: 8–10). This chapter identifies eight basic features of particular interest, though there are others one might mention.

First it is worth noting that philosophers of science hold that a scientific theory must include causal explanation. Following Aristotle, the search for causes has been regarded as one of the core features of scientific theory.[2] While the character of 'causation' itself has been a topic of heated discussion, reinvigorated by David Hume's critique in the eighteenth century, its inclusion in the account of scientific theory has been relatively uncontested. While science must include causal explanation, not all explanation in the natural sciences is causal. When a phenomenon is 'explained' by producing the mathematical law that identifies the quantitative relationship between the variables in question, there need not be an imputation of causation. For example, 'Why did it take three seconds for the apple to fall? Because all objects accelerate at 32 feet per second squared.' Second, when one law is explained by citing a higher-level law, the latter is not offering a 'cause' of the former law. Laws are not causes. So when a law is used as the explanatory factor, the result is not a causal explanation (Cartwright 1983). A causal explanation requires the description of a causal mechanism.

A second feature of the received view is that relationships between variables must be regular, that is, they fit into patterns over time and space. Thus a theory in the natural sciences will include nomic generalisations. A third characteristic is that these generalisations are regular enough to be quantified. Hence the laws or generalisations in scientific theories exhibit mathematical relationships. A fourth feature is that the objects the theory treats and the behaviour they exhibit are believed to be much the same whether they are being observed or not. The entia have existence outside of the minds of the scientific observers, and the scientific observers are observing patterns of behaviour that proceed even when the observers are not observing. Planets and penguins are taken as existing outside of the mind of the scientist.

A fifth feature taken as characteristic of the natural sciences is that all investigators must be able to produce the same conclusions, given the same observations. That is, the differences in conclusions depend entirely on the external world, not on the height, weight, sex, religion, moral creed or other characteristics of the investigator. Hence, the received view of science holds that facts that are observable in the outside world and values that are accepted by the observer are distinct and distinguishable: there is a fact–value distinction and it can be sustained in practice. Science is objective in the sense that results are not tainted by the biases and personality traits of the investigator. Natural scientific enquiry is value-free and any enquiry that is truly scientific will, likewise, be value-free.

Hume and Weber, as is well known, argued for such a distinction. Hume in *A Treatise of Human Understanding* (1965) divided factual 'is-statements' from normative 'ought-statements' and argued that the latter can never be derived from the former. Weber (1949) argued that there must be an 'unconditional separation of facts and the evaluation of those facts'. Factual statements can be proven, which evaluative statements cannot. The two do not mix. Weber admitted that values are involved in the practice of social science in certain ways. For example, the choice of what research to conduct and what research not to conduct is said to reflect the values of the investigator.

A sixth feature of natural science is that the method of gaining knowledge about the external world requires manipulating conditions to create combinations that allow the scientist to distinguish the truth-value of competing hypotheses or theories. This is needed because many crucial observations will not present themselves to the scientist without active manipulation. Experimentation can lead the investigator to learn about the world outside him or her in a way that passive observation cannot. Seventh, the empiricist element in positivism holds that the reports of our senses, under proper conditions, are highly reliable, and science is entitled to its honoured epistemic status because it is a logically systematised method of building on and accounting for these highly reliable reports. This view requires a sharp distinction between observation and theory, because the former must serve as a basis, through a series of strong but imperfect links, to the fallible latter.

The final feature of the received view of science to be mentioned here is that it not only explains the world but allows predictions, at least when enough information is available. Physical theory tells us that if an object is unsupported and is near the surface of the earth, it will accelerate at 32 feet per second squared and will come to rest on the surface of the earth. The moment of its arrival on earth will be predictable, as long as the investigator knows the initial conditions, such as the altitude from which it is dropped. This feature been questioned in the area of the life sciences, notably evolutionary theory, but is still maintained throughout most of the natural sciences. It is particularly valued in the physical sciences, though there are exceptions there, too, such as in some areas of geology.

One may defend naturalism by arguing that all of the features of the natural sciences mentioned above apply to both natural sciences and social sciences, or by arguing that any feature that does not apply to the social sciences does not, upon closer examination, apply to the natural sciences either, i.e., that the received view of the natural sciences was mistaken in its inclusion of that feature of the natural sciences. Partial forms of naturalism are possible, as well, where some selected members of the above list are rejected in the context of the social sciences, but others are defended.

The critique of the positivist account of science

Critiques of positivism are diffuse in nature, stemming from a range of sources; they do not arise from a single doctrine or school of thought. Anti- and post-positivist movements are avowedly pluralistic, offering a wide range of alternatives to positivism. In the social sciences, critics are found emanating from the fields of critical theory (Adorno 2000, Habermas 1987), historical sociology (Skocpol 1992, Tilly 2003), feminism (Elshtain 2003) and postmodernism (Derrida 2002, Heidegger 1969), as well as the scientific/causal critiques (Kuhn 1962, Quine 1990 and their followers). Anti- and post-positivism are represented in IR by scientific realists (Dessler 1989, Patomäki 2002, Wendt 1999 and Wight 1996), feminists (Enloe 2000, Tickner 2001 and Weber 2000), and interpretivist postmoderns (Ashley 1986, Cox 1987, Onuf 1989 and Walker 1993).

There are three sorts of criticisms of positivism that are of relevance here. The first set of criticisms relates to the inadequacies of positivism as an account of science. The second attacks the application of positivism to the social sciences from a non-positivist naturalist perspective. The third set is by authors who reject causal theory altogether and defend constitutive theory in the social sciences, notably interpretivists and supporters of the hermeneutic tradition (HT). This section considers the criticisms of positivism as an account of physical science and the next two sections take up the second and third, respectively.

Several philosophers of science who endorse causal reasoning and the scientific approach launched highly influential criticisms of the positivist account of the sciences. One of the most influential attacks on positivism came from Quine, who criticised what he called two dogmas of empiricism: the distinction between analytic and synthetic statements and the claim that all statements of science are reducible to observation statements.[3] Another attack was the rejection of the observation–theory distinction by means of Hanson's argument of the theory-ladenness of observation, Sellars' 'myth of the given', Scriven's rejection of explanation–prediction symmetry, Kuhn's IP thesis, scientific realists' attacks on positivism's limited ontology, and Quine's and others' attacks on the d-n model of explanation. These criticisms helped erode support for positivism in Anglo-American philosophy of science in the 1960s and 1970s.

It is also worth pointing out that the now-popular resurrection of a form of essentialism in analytic philosophy, which arises from quantified modal logic (which Quine has attacked as excessively Aristotelian) includes an interesting twist on the analytic–synthetic distinction. Quantified modal logic requires that there are propositions that are both empirical or *a posteriori* and at the same time necessarily true. As noted (pp. 29) Sidelle (1989) has offered an interesting argument showing that the necessity is not 'out there' but rather is mind-dependent. More importantly, he has shown that the causal theory of reference – widely accepted by philosophers and endorsed by IR theorists, such as Wendt, who have taken an interest in the matter – requires a notion of 'analyticity' (Sidelle, 1989: 139, 167–9, 198–9).

Quine seems to argue in 'Two Dogmas of Empiricism' (1953) that the principle of radical underdetermination of theory implies that one may hold onto one's favourite theory 'come what may' with regard to new evidence: one may always concede other beliefs and reconcile one's favourite theory with the accepted evidence. Quine's view leads to relativism, since all theories can be reconciled with the evidence by making appropriate adjustments in the body of accepted beliefs. Even more relativist is Quine's view that the given theory and its rivals are not only both supported by the evidence, but are *equally well* supported by it. Lakatos and Feyerabend support this position.[4]

Thomas Kuhn offers another of the most well-known and influential objections to positivism, the IP thesis, attacking the positivists' presumption of an 'objective' basis for testing and comparing theories. IP holds that competing scientific paradigms cannot be compared and that theory choice, in a traditionally understood scientific way, is impossible.[5] Since each paradigm must be understood, according to IP, hermeneutically and as insulated from others, there are no rules or standards *outside* of the competing paradigms which could be taken as 'objective' and employed as a basis for unbiased choice between them. While Kuhn stated such a view, there are places where he offers a less relativistic account (see Smith 1996: 16).

A second argument for IP is that an objective, or intersubjectively valid (held to be the same by all or most observers), comparison of empirical theories would require empirical facts or observations against which the theories or paradigms would be judged. Yet according to this view, no such observations exist, since all observations are influenced by the theories the observer holds. This argument was developed by Hanson (1958). A related argument is offered by Sellars (1963). Others (such as Feyerabend 1978) have since argued that there is no such thing as theory-neutral observation on the basis of which competing theories or paradigms could be judged. (See the useful account of Hunt 1994.)

The d-n model has been attacked in various ways. One traditional source of criticism of the d-n model focuses on its claim of explanation–prediction symmetry. Scriven presents his influential objection (1959), according to which evolutionary biology is able to offer explanations of what has evolved but not predictions of what will evolve. He concludes that explanation and prediction cannot be regarded as entirely alike. Critics also charge that symmetry fails because many events that can be explained after the fact cannot be predicted because the necessary information is impossible to obtain in advance of the event. For example, the rage of a killer cannot be known to be sufficient to lead to the killing before the action is taken, even though after the fact the state of mind of the killer is evident and can be used to provide an explanation (Geertz 1973).

Hempel's view, however, does not require as close a relationship between explanation and prediction; he does not demand that anything that is explainable be predictable. Hempel requires only that the prediction would be formulable, if all of the relevant information were available. Recent criticisms of the symmetry claim include that of Elster (1989: ch. 1), who argues that theoretical indeterminacy is associated with predictive failure but not with explanatory failure. There is an asymmetry also because an indeterminate theory can generate explanatory power, but not predictive power.

One might think that explanation–prediction symmetry is also undermined by Wendt's discussion of multiple realisability (1999: 152–6), according to which macro-level states depend upon micro-level states even though one and the same macro-state may arise from any of a number of distinct (but probably similar) micro-states. But the position Wendt takes is not that there is one single best explanation while at the same time many possible predictions, since he holds that, in this respect, each micro-state would produce a distinct explanation. Nevertheless, he does not think micro-state explanations are always the best available, especially when multiple-realisability is involved (Wendt 1999: 154).

Scientific realism (SR) is, roughly, 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.[6] SR is often viewed as incompatible with instrumentalism and conventionalism. While no detailed argument is offered in support of it here, it is argued below in Chapter 5 that the principle of CS is inescapable but is a safe and non-toxic consequence of a scientific approach to IR and other social sciences. Non-scientific realist theories are often viewed as inconsistent with causal reasoning, though Hitchcock (1992) offers a compelling case for their compatibility.

Authors from a variety of perspectives have criticised the positivist claim that fact-statements and value-statements can be neatly distinguished and that social science can be practised in a value-free fashion. The value-free thesis has been attacked in a number of ways. For example, the Humean claim that one cannot derive 'ought' from 'is' has been criticised by A.N. Prior (1962) with examples from deductive logic. Proposition A is 'George Washington was a Virginian' (factual). Proposition B is 'George Washington was a morally upright person' (evaluative). The proposition 'A or B' is evaluative and follows deductively from the factual statement A alone. A second sort of criticism of the value-free thesis is that while one may separate the evaluative and normative statements A and B based on their semantic content, statements can change from one context to another. The statement 'George Washington was a slave-owner' likewise might be factual in one context and evaluative in another.

It should be noted that evolutionary biology, so far from being a counter-example in another context, has been a source of naturalism in IR theory in two distinct ways. First, many authors have made use of the structure of biological theory, especially evolutionary biology, as a model for IR theory. They have argued that there are many structural similarities between the two. (See especially Modelski and Poznanski and the special issue they edited of *International Studies Quarterly*, 1996 and Bernstein et al. 2000.)

Second, some (e.g., Thayer 2001) have sought to use evolutionary biology as a foundation for the premises of IR theory, arguing that classical realism has rested on weak premises and that the human desire for domination can be understood in Darwinian terms, and then presented as an ultimate cause of the dominance-seeking behaviour of states. Thayer argues that classical realism has been supplanted by neorealism largely because the explanations for, or purported causes of, state behaviour have been unsatisfactory, in light of post-war scientific standards, since they relied on questionable claims about human nature, made by authors like Morgenthau (1949), or metaphysical or religious claims about original sin, made by authors like Niebuhr (1940). The more 'scientifically respectable' third-image versions offered by Waltz and Mearsheimer rely on the notion of 'international anarchy'. But Thayer argues that evolutionary theory provides an even more scientifically respectable and intellectually secure ultimate reason for states' behaviour than does 'anarchy'.

Non-naturalism: scientific-causal critique of positivist social science

Naturalists assert that the parallel between the natural and social sciences can be sustained. One may attack the purported parallel by arguing that any of the above features of natural science do not properly apply to the social sciences. Some authors in the inside and the outside traditions have challenged the social science applicability of all of the features cited above.

All of the features cited above have been challenged by scholars in the outside or the inside tradition (some have been challenged by both) in terms of their applicability to the social sciences. Several of these are the subject of later chapters, such as empirical evidence and the observation–theory distinction in Chapter 3, the nature of social causation in Chapter 4 and the question of prediction in Chapter 5. This section considers criticisms of naturalism coming from the outside tradition, particularly those based on the applicability of the natural-science features of experimentation, fact–value distinction and predictiveness. The next section examines criticisms arising from the inside tradition.

With regard to the fifth feature of the received view of science above, it is worth noting the criticisms of naturalism according to which the social sciences are seen to be inherently evaluative in a way that the natural sciences are not. Non-naturalists of this sort clash with those positivists who defend the notion of the possibility of a value-free social science. They oppose those who argue that the natural sciences inherently must admit value-judgements, if in no other way than in making choices about which phenomena to study (Schutz 1967).

There are important attacks on the applicability of the sixth property of natural science noted above, the role of experimentation. Some of the most interesting work done in the past few decades in the philosophy of natural science focuses on experimentation and laboratory behaviour of scientists. It has been used by figures as diverse as Hacking and Knorr-Cetina. One of the important hallmarks of natural science is the replicability of results performed in laboratories. What the investigator seeks to do in the laboratory is not just to determine what patterns emerge from human manipulation of nature but how nature behaves when humans are not interacting with it, which often requires complex experimental design. The notion of experimental or quasi-experimental design in the social sciences can never hope to accomplish this, since human intervention is the subject matter of the social sciences. Thus it is impossible to attempt to investigate how social patterns develop when there is no human interaction with them.

Since most natural sciences rely so heavily on experimentation to produce reliable results, a problem arises extending the naturalist parallel to the social sciences because experimentation, which has limits of feasibility in the natural sciences, has much greater limits in the social sciences. For example, astronomers are more constrained than some natural scientists. They may be constrained in their observations by the power of their instruments or they may desire to examine a certain phenomenon that occurs only very infrequently. But there are more confining boundaries around the social sciences. For example, psychologists must face the constraint that it is immoral to subject humans to certain conditions and environments for testing purposes. In IR it is impossible to construct an international system with, say, four precisely evenly matched great powers, and to observe the differences between that system and a similar system with two, three or five great powers. If historically there are no cases of four exactly evenly matched great powers, then conclusions about the behaviour of such systems will require much more indirect inference than would be unnecessary if experiments were possible. IR and other disciplines, like sociology and economics, thus rely on quasi-experimental designs based on guidelines of the comparative method, sketched out by John Stuart Mill (1974) in the nineteenth century.[7]

The distinction between observation and theory, the seventh feature of natural science, has been attacked by Hanson (1958), Kuhn (1962), Feyerabend (1978) and others. They hold that the act of observing is not highly reliable, it is not objective and what is observed is not entirely independent of the investigator. Rather, some prior theory conditions the investigator to observe different things in a given situation. Some hold that observation seems objective and uncontroversial only because it is conditioned by widely accepted theories – at least within a given society.

If observation is truly theory-laden, however, then it would seem to follow that theories would be self-affirming and always corroborate the accepted theory, creating a circularity. The basis of the argument would imply that this charge applies to the natural sciences as well as the social sciences. However, in the former case, a response is that observations in the natural sciences have at times been used to overthrow accepted theories. This would seem to refute the charge, at least as directed at the natural sciences. The criticism applied to naturalism holds that it fails only in the case of the social sciences. This raises a difficult question. While observations have served to overthrow received theories in the natural sciences and put them to rest, what examples show that this has happened in a similarly conclusive way in the social sciences?

There are also attacks on the predictiveness of social theory, the last feature noted on p. 38 above, which have been formulated in several ways. According to one, prediction is impossible because of the complexity of underlying mechanisms in the social world. Critics argue also that it is not possible to formulate quantifiable laws. Mill argued that the complexity of social phenomena rendered them outside of the realm of Newtonian-style analysis (discussed in Chapter 5 below; see also Mill 1974, Hayek 1973–8, vol. II, ch. 10). Many of these issues are considered in Chapters 3–7.

Non-naturalism: the interpretivist/hermeneutic critique of positivist social science

The position most thoroughly opposed to naturalism, which flatly rejects any application of the scientific approach to the social world, is found in interpretivism and the hermeneutic tradition (HT). These scholars are part of the 'inside' group, noted in Chapter 1. They hold that IR and other social sciences enlighten the investigator by providing a way of interpreting events, just as hermeneutics in literature or art is a way of interpreting works of intentional human production. The HT rejects naturalism's parallel between the study of the social world and the natural world because the former but not the latter is replete with intention and meaning. Because the subject matter of the two is very different, the appropriate method of study and theorising in the social sciences must be radically different from that in the natural sciences, at least on the received view of the latter.

Hermeneutics and the linguistic model

The HT emphasises that social enquiry properly conducted provides the investigator with understanding through a process much more akin to the study of language than the study of natural phenomena. The HT holds that written texts are decoded hermeneutically (Von Wright 1971: 4–5) and a similar decoding is what is needed as social theorists analyse social institutions and relations. Rules of grammar, symbols, intentions, meanings and signification are the elements of the study of IR. Russia's announcement during the Kosovo crisis that it would redirect its inter-continental ballistic missiles (ICBMs) back onto American targets was a signal to the US of Moscow's disapproval of and anger over the bombing of Yugoslavia; it did not have any material effect on the nuclear balance. Scholars in the HT wish to learn the meaning of this symbolic action, which requires producing a system of meaning in the framework of which an answer becomes possible.

Many arguments, drawn from the HT and elsewhere, attack the claim that the behaviour of the objects of study proceeds independently of observations of the objects or indeed independently of whether there is any science concerning their behaviour at all. That is, they deny the fourth feature of the eight features cited above (pp. 37–9). These attacks on naturalism emphasise the *reflexivity* of the social world, since human beings are the objects of study in the social sciences and they may be influenced by the results of social scientific investigations. Thus theories of the social world affect the world studied in the social sciences.

There are many examples of how social science theories have affected the course of world politics. American foreign policy was surely different than it would otherwise have been when Wilson followed many of the precepts defended by Kant in his theory of an international federation of free states and when President Richard Nixon and Secretary of State Kissinger followed the balance of power principles of classical realists in formulating their Middle East policy. The idea of self-fulfilling or self-defeating prophecies are special cases of this feature. Non-naturalists argue that there is no analogy, including the uncertainty principle, for this in the study of chemical reactions or the behaviour of planets.[8]

Carr, a classical realist, holds that the natural scientist seeks to offer a 'true report' about facts that 'exist independently of what anyone thinks about them. In the political sciences … there are no such facts' (1964: 3). And he says, 'Every political judgment helps to modify the facts on which it is passed. Political thought is itself a form of political action' (1964: 5). Wendt, a scientific realist, sees culture as creating self-fulfilling prophecies (1999: 186–7). Shared ideas about oneself and others are a prerequisite for culture. These shared beliefs give rise to continuity of action; they are continually reinforced. False beliefs about others that are inconsistent with their beliefs about themselves will create conflict and the beliefs will eventually be falsified. Shared beliefs about the existence of the Cold War helped define the identities and interests of the US and the USSR, upon which the superpowers acted. Culture 'tends to reproduce itself, and indeed must do so if it is to be culture at all' (Wendt 1999: 187).

A related problem is that in the social sciences there is some debate about the appropriateness of 'thin description' versus 'thick description', often centred on Geertz's (1973) application of Gilbert Ryle's (1949) important distinction. Thin description includes only observable behaviour of the objects of study; thick description will involve significance and meaning and will use the vocabulary of the subjects. In the natural sciences there is no such debate between the appropriateness of thin versus thick description. It should also be noted here that the fact of self-understanding on the part of the objects of study in the social sciences is no guarantee of the correctness of those self-understandings. Freud, Marx and others have argued that self-delusion and false consciousness prevent the investigator from taking as the final word the objects' self-understandings, their own motivations and their understandings of the structures in which they operate.

The rejection of the elements of the natural science method

HT authors reject various features of positivist scientific method, including causality, nomic generalisations, value-free social theory, objectivity (both in the sense of value-free social science and 'investigator-independence' of the phenomena) and prediction. First to be considered is the HT's rejection of causality and nomic regularities in the social world. The HT rejects causal explanation in several ways. The linear process from one cause to another cannot be identified in the social sciences because there is no linearity; there is an inescapable hermeneutic circle. The linear, thoroughly causal and objective methods of the natural sciences are not appropriate in social enquiry because the objects of study are neither linear, causally connected nor objective but rather are inherently circular, non-causal (that is, constitutive) and subjective.

The HT offers several ways of rejecting a causal attempt to explain observed behaviour, namely, by noting the importance of intentions, of rules and the circularity of context. Suppose someone sees two people conversing and, in walking past them, hears only one word clearly uttered: 'smoking'. He or she may draw conclusions regarding the subject matter of the conversation only if he or she has beliefs about the rules that govern the utterance of the sounds that were heard. The rules will, of course, differ depending on whether the person heard the conversation in Bismark or Barcelona. In North Dakota it is likely that the people were speaking English, in which case they were probably taking about the activity of inhaling particles from smouldering tobacco-filled cylinders placed in their mouths. If the conversation were overheard in Spain, it would be highly probable that the people were speaking Spanish and were taking about formal attire, since 'smoking' in Spanish is the English equivalent of 'tuxedo'.

The rules that govern the action of uttering the sounds heard in the conversation make that conversation what it is – either a conversation about cigarettes or tuxedos. That is, the rules help to constitute the thing itself. The Spanish word 'smoking' has the meaning of the English word 'tuxedo' only because of the rules in virtue of which the sounds constitute that word. Thus rules of language are constitutive rules. In IR, likewise, the action of signing a treaty is only a treaty-signing because of the rules that make it so. Otherwise it would simply be spreading ink on parchment.

The rules in no way provide causal force, rather, they help constitute the action. The treaty-signing is a treaty-signing because of the constitutive rules of international relations and international law. Constitutive rules do for interpretivists some of the work of answering questions of 'how an action came about' that causal regularities do for positivists. As Wendt (1999) puts it, they answer 'how possible?' rather than 'why?' questions. They manifest one aspect of the circularity that social analysis must confront but they do not constitute causes.

The hermeneutic circle can be seen in a variety ways, such as the circularity of intentions, rules and contexts. In the social sciences, unlike in the natural sciences, the actors have intentions. So to provide an intelligible description of behaviour requires that the intentions of the agents be taken into account. It is not possible to identify a social action without getting 'inside' the event. The intentions of the actors are crucial to making the action what it is, and the system of rules in which the action is performed are likewise 'constitutive' of the action in a way similar to rules.[9]

Suppose someone is observed carrying a knife and is seen rushing towards then plunging the knife into the fallen victim of a recent shooting. HT authors note that in order to determine whether the action was a homicidal act designed to guarantee that the victim would not recover or an act of mercy trying extract the bullet from the victim's chest, one needs to know the agent's intention. By looking for the motive one is not looking for something separate from the action. The intention of the actor is part of what constitutes that action, according to the HT. The intention is not thereby a cause of the agent's action, it is a part of that action; it is what makes it a rescue rather than a homicide.

When one is studying rescue attempts, one does not follow the method of relating independent to dependent variable; that is, one does not first isolate motivations or reasons for agents acting to rescue others, then isolate cases of rescue attempts and then try to formulate some sort of relationship between these two independent things. This procedure is unavailable because the investigator would not have been able to identify that particular action at that particular time as a rescue and not an attack if he or she did not already have a belief about the actor's intention. The two things are not independent of one another. The intention is part of what makes the physical action a rescue attempt.

A second way to see the circularity is that the HT shows that without a system of rules it is likewise impossible to identify or classify the action. The action under examination is not ontically distinct from, and is placed into, a framework of rules that give the action meaning and creates a circularity. In a constitutive relationship the rule is a part of what the action is, i.e., part of the action's identity. The rule must be understood if the action is to be properly identified and classified.

For example, when White moves the rook diagonally to threaten Black's king, the action is not counted as a 'move', much less a 'check', because it violates the rules of chess. Likewise, when the blade of a sabre fencer makes contact with an opponent's leg it is not a '*touché*', since the leg is not part of the defined target area in sabre fencing. When a judge adjourns court until 2 p.m. to pronounce sentence on the guilty party and then at the appointed time finds him or herself alone in a lounge, facing four empty martini glasses, banging the gavel and muttering 'Ten years in Levenworth', one cannot conclude that 'sentence has been pronounced', even though the judge performed the same physical actions that he or she would have done in the courtroom, where those same actions would have constituted pronouncing sentence. A description of a '*touché*', a 'check' or 'pronouncing sentence', in each case, will include reference to the relevant rules. This is not parallel to anything in the study of the natural sciences. Rather, it more closely parallels the study of language, where the rules of semantics, grammar and syntax constitute words and phrases, as was evident in the example of the passer-by who overhears a single word of a conversation.

The context is also essential to an understanding of the phenomenon. A focus on the context further illuminates the hermeneutic circle and lack of linearity. In the HT, a description of a social action is not based on 'brute facts' or objective observation-statements; such do not exist, at least for social phenomena. The attempt to make an action intelligible requires interpretation. But the process is circular because the action must be interpreted in some framework of meaning and the choice of the framework will depend on the observed actions. There are many contexts in which any given physical movement or action occurs. A description of the action requires a choice of one (or many, but certainly not all) of those contexts.

For example, when UNSCOM released its December 1998 report on Iraqi compliance with post-Gulf War restrictions, President Clinton ordered an air strike on Iraq. In that case he was operating in many contexts, including those of: a world leader, a president facing an impeachment vote, a husband and father with problems in his relationship with his family. Was the action a punishment of Iraq for poor compliance; an attempt to delay the US Congress on the day of a scheduled impeachment vote; or a reminder to wife and daughter that his actions, whatever they may have been, were performed under pressures and burdens of a magnitude unique in the world?

There were many days when President Clinton did not order raids on Iraq. One makes sense of his decision on 15 December to do so and not on other days by fitting the action into a larger pattern of actions, which can be accomplished only by reference to purposes beyond the observable act itself. It is thus a matter of interpretation which context (world leader, embattled politician, troubled husband and father) one uses to make sense of the action. The circle is completed by the investigator's need to choose other actions (described in one of many possible contexts) in order to see which context makes the most sense in the present case.

With respect to the fifth feature of the received view of science, it is important to note that HT authors reject the fact–value distinction. Some use arguments like those of scientific/outside critics, such as Hanson (1958) and Sellars (1963), discussed above. According to the fact–value distinction and the claim that social science should be value-free, the results of social enquiry are independent of the particularities of the investigator. While interpretivists and scholars in the HT reject the fact–value distinction, there is more diversity of perspective in the fourth feature, namely, observer-independence. Many, but not all, HT scholars hold that while values play a role in theory selection, different scholars can agree on the proper theory choice because of the possibility of value-agreement.

Interpretations of various scholars, based on different frameworks of meaning, for example, may compete with one another. A key question is how one is to judge between them. The dominant answer to this question, taking its inspiration from critical theory (e.g., Adorno 2000, Gramsci 1994 and Horkheimer 1993), is known as 'critical interpretivism'. According to this view, while theories in the human sciences do not contribute to some fictitious idea of objective reality, they nevertheless must serve the human need of emancipation. The value they emphasise, human emancipation, is part of proper meta-theory and must be exhibited by any theory, if it may be regarded as acceptable, that account for and give meaning to events (or 'facts' about events). Critical interpretivism holds that there are meta-theoretical bases for a criterion to choose the best interpretation.

Some interpretivist scholars (e.g., Lyotard 1993) disagree that such a criterion is available. These 'radical interpretivist' scholars deny that there are clear grounds for adjudicating between contending interpretations. They argue that different interpretive accounts or 'meta-narratives', will simply remain as alternatives to one another. Emancipation, as the critical theorists stress, is not given outside of the theory; it is the emancipation of some specific group or individual, not universal emancipation as critical interpretivists seem to believe. Radical interpretivists argue that the critical theorist's reliance on a criterion of emancipation is as ill-conceived as the positivist's reliance on the criteria of verifiability or corroboration.

HT authors reject prediction in the social sciences. As noted, they do not believe that the social sciences admit of causal or nomic regularities. There are several exceptional cases where authors who adopt some of the insights of the HT, nevertheless, hold that nomic generalisations and causal claims are possible. In the philosophy of social science Bhaskar (1975), Bohman (1993) and Harré (1986) are major figures who make this claim; in IR Wendt (1999), Dessler (1989), Patomäki (2002) and several others who adopt Bhaskar's critical realism also maintain that causal generalisations are possible in IR.

In order for prediction to be possible, some regularity must connect present events with future states in a universal, or at least probabilistic, way. Without any such regularities it is not possible to tie the expected or predicted future state into actions taken in the present. So the predictiveness of social theory has no basis in the HT account of social science theory.

In IR the most prominent critical theorists reject prediction. Cox says bluntly, 'It is impossible to predict the future' (1987). Few IR scholars outside the HT remark on this feature of constructivist critical theory. But, for policy-making purposes, the rejection of prediction is quite debilitating for proponents of HT. Chapter 1 argued that policy-making requires prediction, and these scholars deny that possibility. Mearsheimer (1994–5: 43–4) is one of the few 'outside' theorists to note this defect in the inside position.

Mearsheimer, however, does not go quite far enough in his critique. Cox, Ashley and other critical theorists are concerned, as most IR theorists are, with the need to create a better world. Mearsheimer notes that the policies recommended by constructivist critical theorists may bring about changes other than those they desire (creating a more internationally cooperative world without the conflictual influences of realism as the hegemonic discourse). Mearsheimer raises as an example the possibility that ridding the world of realist hegemonic discourse may lead to its replacement with fascist hegemonic discourse, which could create a more oppressive world rather than a less oppressive world. But the problem of unpredictability seems even deeper, since it is entirely possible that the institution-friendly discourse might replace realism as dominant, and there may be no change whatsoever in real world politics. If scholars reject causal (probabilistic) connections between events, states of affairs, or event-types, then there is no reason to believe that any specific change will lead to a particular effect.

There is a somewhat milder form of HT non-naturalism, which does not take a moderate line regarding the core HT tenets. It adopts the core tenets so thoroughly that some of them can be applied to the natural sciences, as well. In comparing science to religion, Peter Winch (1958) argues that the two are much more alike than commonly recognised, since each sets out a system of self-warranting beliefs and principles to order experience. In neither case is there an external or objective standard for determining what is real or rational. Each relies on criteria for what is real that are internal to the theory, and in this sense each is self-warranting. Latour (1987) and Knorr-Cetina et al. (1993) have also argued that circularity is not confined to the social sciences but extends to the natural sciences. They hold that the traditional understanding of the natural sciences is flawed because it does not recognise that it is a thoroughly social activity. The correct understanding of the natural sciences is achieved by likening them to the hermeneuticist's understanding of the social sciences.

Inside, outside and the hypothetico-deductive method

One last point is worth adding here regarding the relationship of the hermeneutic/inside and the scientific/outside approaches to the social sciences, and in particular their shared methods. Despite the fact that the HT rejects the application of the scientific method to the social world, advocates of the HT, nevertheless, share with scientific social theory a number of methods and forms of inference. To take the most obvious example, both use deduction as a method. Proponents of the HT along with theorists in the scientific vein recognise that if contradictory propositions are included in a theory, the theory requires revision. Both approaches make use of probabilisitic arguments. And, what is much less frequently acknowledged, both make use of the h-d method. The hermeneutic process of analysis, e.g., as described by Von Wright (1971) on p. 44, has very broad scope. Pollins (forthcoming: 16) says, 'Indeed Dilthey considered "any manifestation of the human spirit" to be fair game for hermeneutics'. Pollins adds that in the hermeneutic process:

> Fragments of the text are placed in their larger context, understood and given meaning through their location in that context, which itself comes to be reinterpreted as new meanings and understandings are attached to its component fragments. This is the hermeneutic circle – the continual reinterpretation of parts and whole in terms of each other.
>
> (Pollins forthcoming: 16)

The analyst, whether coming from the outside or the inside tradition, takes a conjecture or hypothesis and subjects it to scrutiny. There are competing hypotheses to account for what is known (the movements of electrons, political behaviour or the meanings of words on a printed page). The hermeneutic analyst will employ reasoning of the following sort:

> 'If my interpretation or hypothesis is correct, then I should observe…' When the observations fit with the hypothesis there is further ground for accepting that hypothesis and when they conflict with it or other hypotheses, then the observations provide grounds for rejecting it or them. Dagfinn Føllesdal (1979: 319) says … *the hermeneutic method is the hypothetico-deductive method applied to meaningful texts* [Emphasis in original].
>
> (Pollins forthcoming: 17)

Ted Hopf, a supporter of the hermeneutic approach to IR, describes one aspect of that method as applied to the social sciences.

> If an interpretivist believes that a particular intersubjective understanding implies a particular outcome, whether an action or a cognitive apprehension, then she should a. demonstrate that this relationship holds in her case; b. show that this particular understanding implies additional outcomes in domains unrelated to the particular outcome of interest; c. show that an intersubjective understanding different from the one being investigated in fact implies outcomes different from the one being investigated.
>
> (Hopf forthcoming: 30).

This is a clear HT example of h-d method, as described above on p. 35.

Naturalism and critical realism

In the 1970s, Roy Bhaskar (1975, 1978) developed a distinctive argument for naturalism, which in the years that followed has gained a number of adherents, including in the field of IR (see Chernoff 2002). While Bhaskar endorses naturalism and the claim of a parallel between the natural and social sciences, he rejects many of the aspects of the traditional account of science, not only as they apply to the social sciences but as they apply to the natural sciences as well. In this way the parallel between the two may be endorsed. Bhaskar seeks to replace the positivist–empiricist account of natural scientific knowledge, and especially its account of the objects of scientific study, with a critical realist account. What is relevant for present purposes is that, taking the objects of study as 'real', Bhaskar argues that a stronger parallel between social and natural sciences is possible than when both are understood along empiricist lines – as doubting the status of the referents of theoretical terms.

Bhaskar holds that the basic divide between scientific theory and hermeneutics can be bridged by an account that allows the scientific/causal theory to subsume hermeneutics. Bhaskar recognises the importance of meanings and interpretation but he contends that scientific-style theory can 'explain' these meanings and understandings (Bhaskar 1978; Smith 1996: 27).

According to the HT, meanings and meaningful structures fundamentally differentiate the social sciences from the natural sciences. Most authors who endorse the HT conclude from this that there is no hope for naturalism. Contrary to the HT and Bhasker, on the received naturalist view of science, the social and natural sciences require certain practices that produce knowledge, such as verification, corroboration and/or falsification. But this fails for the social sciences, since these practices cannot be applied to meaningful structures. The empiricist–positivist approach to science thus shows that if one finds the social world inherently composed of meanings and meaningful structures, there can be no strong parallel between the natural and social sciences. Bhaskar rejects the positivist view of natural science and along with it he rejects the restrictive criteria for theoretically acceptable entities. Bhaskar thus believes that there is room within a scientific-style ontology for the sorts of meaningful structures that the HT emphasises. These are 'real structures' according to Bhaskar and thus are capable of fitting in to a critical realist understanding of science and of standing in causal relationships.

The HT, critical theory and SR all attack the positivist-based 'received view' of the nature and methods of science. But while the HT and critical theory attack positivism in a way that undercuts naturalism, SR's attack on positivism supports a recast version of naturalism. According to SR both the natural and social sciences are concerned with identifying structures and explanatory mechanisms (Keat 1981).

For critical realism, the key distinction between the natural and social sciences is that the former is seen as operating in a closed system, or a system that can be experimentally closed, and the latter in an irremediably open system – one in which it is impossible to isolate a limited number of forces, even artificially in an experimental setting. (Bhaskar distinguishes the real from the actual, and part of the actual natural world can be closed.) Bhaskar acknowledges that one kind of mechanism, whether in the natural or social sciences, may be explained in terms of another mechanism, but it does not follow that the former may be reduced to the latter. The human sciences cannot be reduced to material forces. There are, though, some problems common to SR and critical realism.

The most extensive applications of critical realism to the IR theory are Patomäki's recent *After International Relations* (2002) and Patomäki and Wight's recent paper 'After Postpositivism' (2000). While Patomäki follows Bhaskar quite closely, there are points on which he prefers Harré's version of critical realism. Patomäki emphasises that critical realism is based on three central elements, ontological realism, epistemological relativism and judgemental rationalism. There are several important difficulties in critical realism, which it shares with SR, discussed in the next section.

Naturalism, SR and anti-SR

The doctrine of SR has much in common with the more recent critical realism of Bhaskar and his followers. The doctrine of SR focuses on the role of theoretical terms, or those scientific terms that do not refer to observable entities. Philosophers of science have long debated the ontic status of theoretical entities. IR scholars have recently taken a keen interest in it, in the hope that a new view of the matter will help solve some long-standing problems in IR.

Arguments and motivations for SR

Various theories of IR discuss balances of power, cooperative regimes, systems, class conflict, etc. Do these terms refer to things that must be regarded as 'real'? Are there really 'balances' or do we merely imagine them or act *as if* they exist? Are theories of IR about people, states, power, interests, regimes, systems, all of the above or none of the above? Scientific realists and critical realists believe that these questions should be asked at the outset of theoretical enquiry and, for reasons described below, should supplant epistemological questions as the primary foundational steps for an international theory.

Arguments against SR and critical realism applicable to IR and the social sciences are offered elsewhere (Chernoff 2002). Due to space limitations the main conclusions are only summarised here. While common-sense realism, according to which observable entities are accepted as real is only rarely questioned, there is no comparable justification for admitting theoretical entities into one's ontology – especially in the social sciences. The book's causal conventionalism does not endorse the logical empiricist view that such entia do not exist, or the logical positivist view that talk of such entia is meaningless. Talk about them is meaningful and they may exist. But neither does this book accept SR or critical realism. There may well be such entities, but rational grounds are weak for the acceptance of neutrinos, and they are all but absent for admitting international regimes and power balances into one's ontology. This book defends causal conventionalism (CC), according to which causal mechanisms in the social science should be sought and can help provide foundations for prediction, while the theoretical entities referred to in the theoretical statements are not accepted as real by the investigator (see Chapters 4 and 5). Acceptance of these entities enters into one's corpus of knowledge and adds a considerable source of potential error that confers no commensurate advantages to the investigator in guiding policy or in carrying out the goals of building a body of knowledge, a science or a discipline. At the same time, CC does not deny the existence of the entia referred to in theoretical statements.

In the social sciences it is much more likely that laws or behavioural regularities observed today will survive into the next generation of theories than it is likely that the theoretical entities posited by current theories will survive. Any desire to hold onto theories because of the acceptance of the posited theoretical entities is unfounded, since their *only* justification is the theory itself. They provide no other benefits to the theory.

Adoption of SR, according to Dessler (1989), Patomäki (2002), Wendt (1987, 1999) and others, paves the way for major advancements in IR that move the field beyond classical and Waltzian realism. This work has received much positive attention in IR (e.g., Smith 1996: 37 and Wight 1996). Among the most widely cited of recent writings on SR are those of Hilary Putnam of the 1960s and early 1970s (before attenuating his views). According to Putnam:

> [a] realist (with respect to a given theory or discourse) holds that (1) the sentences of that theory or discourse are true or false; and (2) that what

> makes them true or false is something *external* – that is to say, not (in general) our sense data, actual or potential, or the structure of our minds, or our language, etc.
>
> (1975: 69–70, see also Boyd 1973)

Scientific theories contain observational terms and theoretical terms, which refer to unobservables (though there may be some difficult borderline cases). The statements involving theoretical terms must somehow be linked to experimental or observation statements. Some sort of correspondence rules are to serve this purpose. For scientific realists these rules are dependent upon a model for the theory that will supply an interpretation of both sorts of terms. There are significantly different forms of SR, including inferential realism (abductive inference to the best explanation), fiduciary realism (where credibility is at the core), bivalence realism (according to which all statements are either true or false) and referential realism (according to which all entities in scientific theories have genuine ontic status). Some authors identify SR explicitly – and mistakenly – with the natural-scientific view but one may perfectly well accept naturalism and reject SR.[10]

There are several important difficulties with the doctrine of SR. One relates to its account of theoretical terms (the crux of the doctrine), which critics charge is hopelessly flawed. Scientific realists believe that scientific theories refer to real entities. But since fallibilism accepts the possibility that the best theories will be overthrown in the future and the theoretical entities replaced by others, scientific realists face a serious problem, unless they resort to an infallibilist theory of knowledge. Attempts to address this problem have severely weakened SR (McAllister 1993). Another charge against SR is that the structure of the explanation it offers is formally invalid (Laudan 1981). It charges that SR essentially has to use the principle of inference to the best explanation to justify the principle of inference to the best explanation (Laudan 1981: section 7). This is perhaps 'the most telling' of all criticisms advanced against SR (Kukla 1996: S298).

Motivations for adopting SR

There are, then, clear difficulties that make adoption of SR a far from cost-free choice. One must then ask, why do many scientists and philosophers call themselves 'scientific realists'? Beyond the fact that it seems to capture common sense, two main motivations underlie the desire to endorse SR, according to its proponents. First, SR provides the only account of the progress of science that does not make the success of science a miracle and, second, it is the only account that leaves open the possibility of causal explanation. But there are difficulties with both motivations for SR in the social sciences. The first is troubling because natural-science progress has no counterpart in the social sciences, since the history of the social sciences is so unlike that of the natural sciences. That is, scholars often emphasise the contrast rather than the parallel on this point – the enduring lack of consensus between political realists, liberals, idealists, etc., on most questions of IR theory. The second motivation is irrelevant, since those who wish to retain causal explanation have options other than adoption of SR (see Chapter 4 and Chernoff 2002: 198–9). There is then no compelling motivation to adopt social SR. (Though various versions of SR remain popular among philosophers of science, Putnam, who, as noted, was one of the most respected of scientific realists, eventually abandoned the doctrine.)

Wendt offers two motivations for adoption of SR. The principal argument is that anti-SR foundations rule out various legitimate IR theories by means of *a priori* arguments, which should not be permitted and which SR does not do. The purpose of Wendt's social scientific realist enterprise is to block '*a priori* arguments against engaging in certain types of work, [wherefore, scientific] realism is a condition of possibility for the argument in the rest of this book' (Wendt 1999: 91). His argument is roughly as follows: first, meta-theory should not rule out substantive theories one might otherwise accept. Second, SR has no such effect but anti-SR does. Third, one should, therefore, accept SR as one's meta-theory. Wendt's argument is unsound because both premises are false. The first premise is false because one should have some *a priori* criteria for a properly formed theory, which would at least include requirements like 'internal consistency'. The second is false because SR does require that assumptions be 'realistic' and avowed instrumentalists need not accept this requirement.

Wendt's second argument claims that most IR scholarship tacitly accepts SR, even in cases where theorists identify themselves as empiricists (1999: 47). Wendt argues that scholarship in the field of IR will be more fruitful and more likely to show progress if the foundational debate is shifted from epistemology to ontology because there is, beneath the surface, more agreement among disparate IR theorists on the latter than on the former (1999: 91). Patomäki and Wight (2000) offer a similar argument. Wendt reasons as follows:

1. agreement on foundational questions aids progress in a discipline;
2. most IR scholars accept SR, even if they do not acknowledge it openly;
3. Wendt infers from 2 that there is less disagreement among IR scholars about ontology than about epistemology; therefore
4. a shift from epistemological to ontological foundations aids progress.

This second argument is invalid because premise 3 does not follow from premise 2. Even if IR scholars accept SR, it does not follow that they share ontological views, since SR merely states connections between a theory and ontological commitments. It does not commit IR theorists to any particular ontology – because it does not lead to adoption and any specific substantive theory. Scholars may then disagree about any particular ontology as much as they do about epistemology. Wendt manages to reach his conclusion only by mischaracterising the debates in the philosophy of social science, forcing the choice only between those who accept the reality of social entia and those who reject them, overlooking those who suspend judgement (Chernoff 2002: 204–5).

Thus both of Wendt's arguments motivating SR fail: the first argument is unsound and the second invalid (see Chernoff 2002: 194–9). Wendt correctly rejects the scepticism arising from the principles of radical underdetermination of theory by data and incommensurability of paradigms (1999: 66). But he is incorrect to contend that SR is the surest foundation for the epistemologically pluralist position he advocates. The problems Wendt poses may be solved without the ontological baggage of SR by means of conventionalism, discussed below in Chapters 5–7. (See also Grünbaum 1968, Kyburg 1990b.)

In addition to his arguments motivating SR, Wendt offers arguments attacking anti-SR positions, instrumentalism in particular and empiricism in general. He charges that instrumentalism ignores the matter of whether a theory's assumptions are 'realistic' as long as it produces good predictions. But his response regarding the ontic status of theoretical entities begs the question by assuming SR is true. Second, Wendt says that instrumentalism eliminates some theories from consideration on *a priori* grounds, which is unacceptable. But some *a priori* criteria are entirely appropriate, such as internal consistency and coherence.[11] Wendt argues against empiricism by levelling three charges, that the d-n model was only intended as an ideal, which social scientists lose sight of; that scientists' ability to manipulate the world makes it less 'reasonable' to doubt that the deep structure of the world is better known now than it was in Hume's era; and that subsumption under a law does not properly qualify as an explanation. All three criticisms, similarly, can be shown to miss their marks (evident in Chernoff 2002: 196–9).

Wendt is, consequently, mistaken in saying:

> [t]he primary significance of [scientific] realism for causal theorizing is in cases where lawlike generalizations are not available, either because we are dealing with unique events or because the complexity or openness of the system defies generalization. In these cases the logical empiricist would have to give up on causal explanation; the realist would not.
>
> (Wendt 1999: 82)

While this may be true of anyone who endorses Wendt's narrow notion of 'logical empiricism' (see also Shapiro and Wendt 1992), it is not true if 'empiricism' is taken, as it is by Wendt, to stand for the non-post-positivist alternative to SR. There are anti-social scientific realists like Little (1991) who need not give up on all forms of causal explanation. Wendt's argument for SR collapses because it is built on his mistaken characterisation of the competing philosophical positions and their interrelationships. Wendt is wrong in saying that SR 'is a condition of possibility for the argument in the rest of this book' because of its unique ability to block '*a priori* arguments against engaging in certain types of work' (Wendt 1999: 91). An anti-social scientific realist meta-theory like Little's does it just as well. The 'difference that SR makes' for social science theorising is far from what Wendt claims.[12]

The move to ontology from epistemology

It is important to keep in mind what the purposes of philosophical foundations are. SR, critical realism, and Patomäki and Wight's 'philosophical realism' which entails SR, violate the basic motivation for seeking philosophical foundations, which, as Socrates pointed out, requires first that one must purge oneself of illusion and false belief, and second that one develop a corpus of true beliefs that helps to guide actions in life.

The desire to seek a philosophical foundation for IR does not commit one to 'epistemological foundationalism', namely, the attempt to build a body of knowledge based on a set of indubitable or incorrigible 'foundational' propositions, typically reports about sensory experience. Rather, 'foundational' is used here more broadly to refer to a set of philosophical and methodological principles that purport to solve questions raised by (social) science theory and practice. One may offer an epistemological foundation without embracing 'epistemological foundationalism', as Wendt, who rejects epistemological foundationalism, does when he says (1999: 48), 'I provide the *foundation* for the [scientific] realist claim that states and the state system are real (ontology) and knowable (epistemology), despite being unobservable' [emphasis added]. Others who reject foundational theories of knowledge also seek a 'foundation' for IR theory (e.g., Jørgensen 2001). Almost all of those who reject the doctrine of 'epistemological foundationalism' still agree that the acceptance of some propositions depends logically on the acceptance of others. In this limited way a foundational theory, or metatheory, is useful even for those of us who are, in the philosophical sense, epistemological non-foundationalists.

One must take some risks in developing a corpus of knowledge. Absolute certainty is not a reasonable goal or standard because empirical knowledge is fallible. However, one must seek to avoid introducing the possibility of error without the potential for significant gain. If the danger of error is counterbalanced by gains in knowledge and guidance for people's lives, it would be rational and epistemically justifiable to accept new propositions that raise the chances that one is in error. But introducing the potential error of belief in the existence in positrons, anti-matter and especially social science entities like international institutions or forces of supply and demand, does not offer comparable benefits (as emerges from the acceptance as real of unobservable entities referred to by the prosecution at the witch trials, alluded to in this chapter's epigraph by Arthur Miller). One may carry out the work of science and foreign relations without admitting the danger of these sorts of error.

Naturalism in contemporary IR

Among recent IR theorists naturalism, or something close to it, is widespread, particularly for those who reject constructivist, postmodern and reflectivist schools and the HT.[13] Contemporary IR theorists of the liberal and realist traditions endorse a good deal of naturalism. Even some constructivists part company with the majority of their postmodern and reflectivist colleagues in endorsing some naturalist tenets. This section offers a brief discussion of the most prominent current IR theorists' views.

Contemporary neorealists tend toward naturalism much as their classical realist predecessors did. Waltz is explicit in announcing his naturalism. He frequently treats 'theories' in such a way that IR theory is structurally interchangeable with physical theory. Waltz speaks of 'theory, whether Isaac Newton's or Adam Smith's' (1979: 10). Mearsheimer holds similar views, evident from the way in which, for example, he articulates causal mechanisms of competing theories (1994–5). Waltz and like-minded IR theorists cite the philosophical works of Einstein, Hempel, Popper, Carnap and Putnam to support their view of IR meta-theory, clearly invoking the model of the natural sciences. Much of Jervis's recent systemic explanation of international politics (1997) makes use of the model of biology. Jervis uses biological systems extensively as a model to shed light on the proper ways of conceptualising and studying international systems. The biological model has been increasingly invoked by naturalist IR scholars. A recent special issue of *International Studies Quarterly* (Modelski and Poznanski 1996) explored the application of evolutionary models to the study of world politics.

Contemporary neorealist scholars in IR like Waltz and Mearsheimer tend to focus on war, peace and great power conflict, though some (such as Gilpin and Krasner) have written primarily on political economy. Keohane and many other prominent contemporary neoliberal scholars focus on political economy, though a number focus on security, war and peace, e.g., Doyle (1983a, 1986), Fukuyama (1992), Lebow (1981, 2003), Rummel (1992, 1997) and Russett (1963, 1990). Among the most prominent contemporary liberals, Keohane has taken what should be regarded as a naturalist position, e.g., in *After Hegemony* (1984), although still a somewhat minimal version of naturalism. He sees parallels between IR and natural science but does not push the analogy as far as some other liberals. Keohane says that international cooperation 'is particularly hard, perhaps impossible, to investigate with scientific rigor' (1984: 10). He describes his method as 'interpretive', among other things, and says that his sense of values motivated his enquiry. Nevertheless, he contrasts such values with the form of enquiry that comprises the bulk of that work. He states that his theory may be subjected to fair analysis, if not strict scientific-style testing, by other authors who do not share his normative views (1984: 10). Keohane assents to the standard view that, in IR, the 'social laws' or general 'propositions are at best valid only probabilistically' (2000: 127).

Russett, like other mathematical modellers, implicitly endorses naturalism. He refers often to the study of IR as a 'science' (e.g., Russett and Oneal 2001: 313).

In some of his earliest publications (e.g., 1963) he makes use of an approach that develops mathematical models of relations among states that are analogous to the mathematical models of relations among physical bodies.

Vasquez, as much as any contemporary writer, and certainly as much as any opponent of the tradition of political realism, has self-consciously and carefully considered meta-theoretical questions, basing many of his most important arguments on criteria of theory choice drawn from the work of Popper and Lakatos – both primarily philosophers of physics. The implication of Vasquez is very clearly that the similarities in structure of natural-science and social science theories (and 'research programmes') are genuine enough to allow physical-science methods to guide social science theory appraisal.

Anti-positivists and postmodern writers emphasise the question of what science is and how the study of IR may be viewed as such. One does not have to look far to see the anti-positivists' use of meta-theory. In his oft-cited essay, 'The Poverty of Neorealism', Rick Ashley (1986: 280) is explicitly critical of certain 'metatheoretical commitments' of neorealism. Ashley points out that neorealists' critique of classical realism includes a conception of what counts as a 'science' and that neorealists argue that classical realism does not satisfy that definition (1986: 260). Classical realists do not meet 'modern *scientific* standards' of theory, specifically because they are 'too fuzzy, too slippery, too resistant to consistent operational formulation'. Referring to Waltz (1979: 62–4), Ashley notes that neorealists do not distinguish between subjective and objective aspects of international political life (1986: 261). Ashley (1986: 261) also cites Gilpin (1981: 3), who says that classical realism 'is not well grounded in social theory'. But on the other hand, on the question of naturalism (or what he calls 'positivist' social science), Ashley endorses Giddens's claim that 'there are no particular barriers to the treatment of social conduct as an "object" on a par with objects in the natural world' (Giddens 1974: 4, cited by Ashley 1986: 281).

Another influential reflectivist theorist, R.J.B. Walker, is concerned about the uses of natural science in the development of IR and social theory. According to Walker, 'The conventional distinction between the sciences and the humanities obscures more than it reveals' in the uses of models, analogies and 'images taken from Newtonian mechanics or Darwinian biology' (1993: 97–8). He says, 'In the analysis of world politics, the notion of balance of power itself clearly has an analogical quality' (1993: 98). He notes that 'the logic of scientific explanation has been extended from the sciences of inert matter to encompass patterns of probability in historical practices. But such strategies have always encountered powerful opposition. The historicity of human experience remains deeply problematic' (1993: 101). Walker sees the anti-positivist view of the social sciences as having taken over. He says, 'many scholars concerned with what it means to study social and political life have turned away from the largely positivist accounts of scientific explanation to a much broader area of philosophical debate, one in which the explorations of literary theorists are treated at least as seriously as pre-Kuhnian dogmas of cumulative scientific knowledge' (1993: 83).

Wendt agrees with other critics of positivism about the importance of constitutive relations but he consistently tries to show that social structures have both constitutive and causal effects (1999: 165–71). There are methodological differences between the two – neither can one interview bacteria nor can one understand policy-makers' decisions by examining their cell structure. Still, there are profound similarities that other constructivists and critics of positivism tend to deny. Wendt says, 'Positivists think that natural scientists do not do constitutive theory and so privilege causal theory; post-positivists think social scientists should not do causal theory and so privilege constitutive theory. But in fact, all scientists do both kinds of theory' (Wendt 1999: 77–8). Wendt's naturalism is clear in many passages. For example, 'There is nothing in the intellectual activity required to explain processes of social construction that is epistemologically different than the intellectual activity engaged in by natural scientists' (1999: 372).

Wendt identifies two other characteristics as essential to 'science', whether natural or social, namely, 'publicity of evidence' and 'falsifiability'. In both the natural and social sciences the success of an enquiry 'depends on publicly available evidence and the possibility that its conclusions might in some broad sense be falsified' (1999: 373). If the falsifiablility criterion is not satisfied, then one's work would, at best, be 'a form of art, self-expression, or revelation. But it is not a genuine effort to know the world through "science" (1999: 373).

Conclusions: degrees of naturalism

Naturalism is the view that the methods of the natural sciences should be applied to the social sciences in order to achieve the best possible theories, explanations and predictions. There are many ways in which the parallel between the natural sciences and the social sciences have been drawn. This chapter has identified eight aspects of natural science that philosophers have tried to argue are or are not parallel to the social sciences. Most authors outside of the HT hold some of the eight tenets, though most do not endorse the parallels asserted by all eight.

The question of whether all investigators must be naturalists or non-naturalists is misleading, and seems to presuppose a false dichotomy. It is helpful to phrase the question as Bhaskar does when he begins his book by saying the, 'primal problem of the philosophy of the social sciences is' to determine 'to what extent can society be studied in the same way as nature?' (Bhaskar 1998: 1). Bhaskar here recognises that naturalism should be understood as a matter of degree. The present book defends a version of naturalism, indeed a fairly strong version of naturalism, but does not endorse an unqualified version.

The terms 'naturalism', or the rough synonyms 'scientific essentialism' and 'unified science', are usually not specific enough to detail which of the various characteristics they emphasise. Many authors who discuss naturalism do not identify precisely which aspects of the natural sciences the social sciences must share in order for the philosophical position of naturalism to be sustained. This is especially true in the social science literature. HT authors do not, however, face any severe problems on this score, since they generally reject all forms of naturalism in the social sciences. The only exceptions are those who also reject the traditional understanding of science as it applies to the natural sciences. Their position thus opens up a postmodern form of parallel between the two. In any case, since many authors are not specific about which aspects of the parallel they endorse, it is often hard to say precisely who is and who is not a naturalist. It is left unclear just how far they believe one has to push the analogy with the natural sciences to reach the naturalist threshold. That is, how many of the eight features of the received view must one endorse to be a naturalist? And are there some features that one must accept to be termed a naturalist? There are then many non-HT authors who accept at least some parallels, and they would not seem to fit the definition of a thorough-going naturalist; they would seem to warrant classification as more naturalist than the thoroughly non-naturalist, such as HT authors.

If non-naturalism is the view that there are differences between the natural sciences and the social sciences, then almost everyone is a non-naturalist, since at this stage very few groups of scholars would argue that the two are identical. If non-naturalism is the view that the social sciences are thoroughly and radically different from the natural sciences, that is, that there are no parallels, then, almost everyone is a naturalist. The category of non-naturalism would, as noted, include only interpretivist, HT and postmodern authors, e.g., Ashley (1986), Taylor (1985) and Winch (1958). Since almost all philosophers of social science hold that some combination of features 1–8 cited on pp. 37–9 hold in the social sciences, one might term them 'semi-naturalists'. It would be arbitrary at this point to select a minimum number or specific group of the eight features listed above that one must endorse to be properly classified as a naturalist.

Scholars of IR do not always address all eight of these characteristics. In IR major theorists differ on these eight characteristics. Waltz, who fits closely with the received view, nevertheless disagrees with the claim that international structures are mind-independent. Since Waltz accepts the distinction between observation and theoretical terms, he holds that the latter may or may not exist outside of the theory. However, he remains a naturalist, since he denies that theoretical entities in the natural sciences are mind-independent.

Wendt, usually categorised as a constructivist, attempts to synthesise some of the ideational elements of constructivism with some of the methodology of a more traditional view of science. Wendt (1999: 91) puts it by saying that he endorses a positivist epistemology and post-positivist ontology. Wendt would endorse most of these tenets of science as applicable to the social sciences, including the one Waltz dismisses, that international structures are mind-independent; however he would reject two of the features cited: that theory is value-free and that observation is privileged over interpretation. Pure constructivists would reject all eight.

When Wendt is contrasted with pure constructivists, it becomes evident that he agrees with the latter about the importance of constitutive relations. But Wendt consistently tries to show that social structures have both constitutive and causal effects (1999: e.g., ch. 3). The argument developed below will likewise be a semi-naturalist position, agreeing with the bulk of the traditional views applying to the social sciences, but questioning the two objectivity claims: mind-independence and value-freedom, at least under some conditions.

Human beings do not look much like chimpanzees. There are essential differences between the two species. One worships transcendent deities, has moral codes, theorises about politics, and builds skyscrapers, while the other, as far as we know, does not. It thus came as a surprise to many members of the former species to learn that over 99 per cent of their genetic makeup is identical to that of the latter. Knowledge of human behaviour and institutions has struck many scholars as vastly different from knowledge of the natural world. However, both are forms of systematic empirical enquiry and, as such, share a myriad of features, such as application of the rules of formal logic, statistical methods, descriptive categorisation and observation of phenomena, the h-d method, etc. While many reject the naturalist analogy between social science and natural science, there are various degrees of naturalism between outright rejection and adoption of an identification of the two spheres of knowledge. There are indeed some parallels with the natural sciences. The difficult question is how far they extend.

3 Theory, observation and law

UNDERSHAFT: [your religion] … doesn't fit the facts. Well, scrap it. Scrap it and get one that does fit.

George Bernard Shaw, *Major Barbara*, Act III

JUDGE: And the presence of this deposit is conclusive evidence of firing?

WINBOURNE: I'm a scientist, my lord. I don't know what constitutes conclusive evidence.

Brian Friel, *The Freedom of the City*, Act I

In order for policy-makers to evaluate the merits of available options they must be able to answer specific questions. Can the removal of Saddam Hussein and subsequent US policy moves produce a stable democracy in Iraq? Should the US retain a long-term troop presence? Would such a presence be more likely to promote or undermine a stable, sovereign post-Ba'athist regime? Would a US invasion and presence do more to help US security by removing a dangerous Iraqi dictator or do more to damage US security by creating anti-American sentiment that fuels terrorist attacks on Americans?

The debate among scholars over the best policy towards Iraq has produced analyses replete with law-like generalisations. Byman (2003: 49) generalises about the relationship between democratic transitions and external assistance. The Rand team – drawing on past US attempts at nation-building in Germany, Japan, Somalia, Haiti, Bosnia, Kosovo and Afghanistan – offers generalisations about the workings of the US government, internal developments and success, and the effects of third-party states on success (Dobbins et al. 2003: 69, 107, 221). Hollis also discusses historically imposed limits to valid generalisations concerning, for example, domino-type impact of the first Islamic state in a region to democratise (Hollis 2003: 25–6, 31). Brooks proposes generalisations relating to democratisation and the success of Islamist movements (2002: 616). Metz (2003–4: 29) generalises about the effectiveness of various recruiting methods for terrorist groups. And Barton and Crocker (2003) generalise about specific policy options and their success in promoting stabilisation.

Policy decisions about whether and how to support the new regime in Iraq involve predictions, which may be judged to be justified, rational and credible on the basis of one's theory, which in turn is based on the evidence that supports the theory and criteria of theory choice. For example, the claim that continued US military presence is essential for Iraq to be able to create a stable, liberal democracy requires predictions like 'Post-Saddam Hussein Iraq is likely to revert to an illiberal form of government without a US military presence' and 'A US military presence will not produce so much resentment among Iraqis that they will rebel against the political goals of the occupation'. If a policy-maker whose goal is to promote liberal democracy in Iraq does not accept these or similar predictions, then he or she would not have a basis for supporting a policy of continued large-scale US military presence. One needs evidence to accept (or reject) these predictions. Dobbins et al. (2003) and Byman (2003) attempt to collect and evaluate the sort of evidence that is considered relevant. Evidence will come from observations of similar past cases, primarily cases of intervention and occupation, within the framework of a theory. The theory will have behavioural generalisations or laws derived from the observations.

Contending theories and types of laws

Liberals and idealists can agree with political realists on the 'observation' that, throughout history, international coalitions have been formed at many points. But the different theoretical traditions offer quite different accounts of why they did so. Consider the coalitions against Napoleon, against Germany in each world war, against the USSR after the Second World War and against Iraq in 1990–1. Liberals highlight the role of ideas: those opposing Napoleon shared a monarchical ideology and opposed France's revolutionary ideology. In the case of the coalitions against Germany in the first half of the twentieth century, they comprised varying sorts of liberal democracies, like France, Britain and the US, who shared an ideology of democracy and opposed 'German despotism'. In the 1950s the NATO states were liberal democracies opposed to communism and the same argument was applied in the case of the Gulf War, where Saddam Hussein was portrayed as a dictator who had an insatiable desire for territorial expansion and power, and who was a threat to the more moderate or liberal states of the Middle East and democratic Israel.

Political realists object to the liberals' interpretation of these examples. They note that Russia in 1914 and the USSR in 1941 were on the side opposing Germany but did not fit the description of 'liberal democracy'. NATO included Greece and Turkey during periods of military dictatorship in those states. During those periods NATO did not move to expel or suspend undemocratic Greece or Turkey. And the 1990–1 Gulf War coalition, which comprised over thirty states, included Syria and other non-democracies in the Gulf.

Political scientists who accept this realist view will defend it by advancing a behavioural law like (LB1) *the rise of an expansionist state perceived to be threatening will give rise to a coalition of states allying against it*. France, Germany twice, USSR and Iraq are then said to fill the role of dangerous, rising, expansionist powers in the above 'law' of IR.

The law LB1 identifies a pattern that both realists and liberals may equally acknowledge. Realists have explanations for the aggressive or competitive nature of the state, which accounts for why the increasing power of one state is treated by others' collective action. They might invoke the *explanatory* generalisation (LE1a) *the expansionist nature of states leads less powerful states to work together to combat the most powerful in the region or system*. Kantian liberals offer a different explanation for this pattern, namely, that non-democratic states use more of their available resources to develop military might and such militarisation threatens others. One example would be the difference in the proportion of GDP devoted to military expenditure in the USSR versus the US, Nazi Germany versus Great Britain or Iraq versus Kuwait. Liberals would suggest an explanatory law along the following lines (LE1b) *states that share values and/or an ideology will work together to prevent states with an opposing ideology from dominating the system*.

Thus realists and liberals might agree that the behavioural pattern identified in LB1 represents the truth about the international system. But realists and liberals develop contrasting theories, each with their own explanation, to account for a commonly held belief about observable behaviour, LB1. One's desire to know which theory, or explanatory law, such as LE1a or LE1b, best accounts for LB1 raises several questions which occupy this chapter. First, how does one go about making a decision as to which theory to accept as the superior theory, that is, what are the proper criteria for theory choice? A second question stems from the attempt by many theorists to separate behavioural laws like LB1 from explanatory laws like LE1a and LE1b and the theories of which they are a part. Can such a separation be cogently effected?

The distinction suggested here between behavioural and explanatory laws draws on the traditional empiricist distinction between 'observation' and 'theory'. Carnap (1966: 225–31), for example, distinguishes 'empirical' from 'theoretical' laws. However, philosophers of science of a much less empiricist bent offer something similar, such as Cartwright's distinction between 'phenomenological' and 'fundamental' laws and Little's distinction between 'phenomenal' and 'governing' laws (and what others sometimes call 'experimental' and 'fundamental' laws).[1]

How one goes about answering the question of the connection between behaviour and explanation (or observation and theory) will have an impact on one's notion of 'cause' in physical and social enquiry, which is considered in the next chapter. For example, some of those (e.g., empiricists) who draw a sharp distinction between 'observation' and 'theory' use it to develop an account of scientific reasoning that eschews talk of 'causality' altogether. The present question is how one should formulate generalisations and how one should explain them (if at all). Are laws like LB1 the only legitimate type of law?

A central problem of methodology is then to determine how one develops a theory to account for patterns of the sort identified in LB1. The investigator seeks some more fundamental generalisation to explain the pattern. Were the coalitions against Napoleon, the central powers, the axis powers, the USSR and

Iraq formed by a shared idea that democratic values should be spread, as LE1b says, or were they perhaps formed as a result of a shared fear for the allies' safety or their territorial integrity, as LE1a would suggest, in which case it may be asserted that states always balance against power? How can a theorist choose the correct theory and how will this aid in the making of foreign policy?

Observation terms and theoretical terms

The distinction

Traditional empiricist methodology, as is well known, draws a sharp distinction between 'observation' and 'theory' and between the language of 'observation' and the language of 'theory'. All terms used in a scientific language are classified either as observation or theoretical terms. One version of the distinction holds that observation statements are inherently descriptive while theoretical statements are explanatory (Alexander 1963). On this view the behavioural law LB1 above falls into the 'observation' category. There are good reasons to maintain the distinction between 'observation' and 'theoretical' terms, since it aids in the construction of a foundational structure for scientific reasoning. But the distinction must be abandoned if philosophical arguments against it outweigh those that exculpate it. Whether this is the case is the subject of this and the next section.

Kyburg (1977: 92) develops a distinction between observation and theory in which he says that we regard some 'terms… as "observational," in the sense that we can come to accept in our body of knowledge sentences involving those terms on the basis of what happens to us'. One of the interesting and important features claimed for observational laws is that they are true regardless of the theory one accepts. Those who support the observation–theory distinction agree with scientific realists on the question of whether laws bear truth values. The laws are seen as invariant in truth value and, by implication, invariant in meaning. The ideal gas law is a standard example. Whether one views gases as amounting to continuous fluids or as bouncing molecules, it remains true that pv = nrt. In IR there is hegemonic stability theory which includes the behavioural law (LB2) *public goods arise in a hegemonic system*. There are liberal theoretical laws that seek to explain it, which include that the hegemon is, in the parlance of public goods theory, a 'privileged group' (of one) and is willing to supply public goods, like order. On the realist approach, a theoretical law to explain LB2 is based on the view that the powerful coerce those they have the power to coerce and a hegemon, especially, may coerce others into doing what it desires.[2]

Another example of an observational law might be that (LB3) *democratic states do not go to war with one another*. One account, the structural model, holds that (LE3a) *democracies have more cumbersome decision-making procedures than non-democracies* and so, when crises arise, democracies are more likely to get through them without attacking simply because it takes so much time to build the necessary political support and then initiate the attack (Maoz and Russett 1993). The crisis is much more likely to have abated by the time this process is completed. A second account, the normative model, holds that (LE3b) *the internal peaceful norms of democracies are externalised in international conflict*, especially as other democracies are viewed as deserving of rights and considerations akin to one's own citizens (Maoz and Russett 1993). These two accounts provide the basis for explanatory laws that differ from one another, while both still manage to account for the behavioural regularity captured in LB3.[3]

Objections to the distinction

There are several well-known lines of objection to the observation–theory distinction. One is that the distinction is hopelessly confused, as evidenced by the different ways the terms are used in different fields. For example, scientists talk of observing the flow of electrical current in the ammeter or of observing the path of an electron in a cloud chamber. But philosophers do not regard these as observable. Second, the terms are argued to be relative to specific cases. Different scientists will invoke different standards under different circumstances or for different purposes. What is taken as observable in one set of experiments is taken as unobservable or theoretical in another. And third, by far the most widely accepted criticism of the distinction is that 'observation' and 'theory' are never conceptually distinct; that is, observations are embedded in theories or are 'theory-laden'. This last argument claims that different investigators apparently looking at the same phenomena actually see different things. Observation is thus dependent upon and embedded in the theory that one accepts. Most contemporary IR theorists who explicitly discuss this issue seem to endorse this principle.

The third objection, proffered by Hanson (1958) and Sellars (1956, 1965), states that two investigators who endorse different theories (or disciplinary matrices) observe different 'things' when they survey the same visual, auditory, olfactory or tactile field. Their different conceptual frameworks lead them to organise the mass of stimuli data in different ways. In a non-scientific context this might occur when three people from different cultures look at a tree and one apprehends a carbon-based life form, one a source of firewood and the other a resting spirit. At the level of scientific theory, some have argued that when Galileo looked at the sun he saw an object about which earth revolved, whereas when Aristotle looked at the sun he saw an object that revolved around the earth, from which it follows that they were, in a fundamental sense, looking at different things.

The problem, one might argue, is at least as severe in the social sciences, since the social entia with which investigators deal are abstract structures. For example, the most fundamental terms or building blocks of IR theory include 'state', 'war' and 'institution', none of which enjoys the benefits of physical existence. The notion of 'human beings as agents' helps here, since humans are corporeal, but IR theorists are reluctant to confine themselves to an ontology of observables. For example, the ontology of IR includes at least states and wars. However, to cite one example of the problem, if one defines ‘war’ as ‘militarised inter-state disputes involving at least 1,000 battle deaths’, then one has to define ‘battle death’, which requires (what turns out to be) a non-observational distinction between battle-deaths and deaths that occur in other circumstances.

Among IR theorists, the third criticism seems to be almost universally accepted. Wendt takes an explicit position on the question of theory and observation, agreeing with Waltz. Wendt (1999: 370) says, ‘All observation is theory-laden, dependent on background ideas, generally taken as given or unproblematic, about what kinds of things there are and how they are structured’ and that, ‘theory to some extent constructs its own facts’ (1999: 58). Theory-language may differ from observation-language in the degree to which it presupposes background beliefs, but it does not differ in kind (1999: 62).

Waltz deals specifically with the relation of observation to theory. He takes a clear position on this question, endorsing the view that all observation is theory-laden (1997: 913–14). He assents to the widely held view in stating, ‘Scientists and philosophers of science refer to facts as being “theory laden” and to theory and fact as being “interdependent” (1997: 913). Ruggie (1986) finds that one of the faults in Waltz’s structural realism is the unobservable nature of his ordering principle. International systems are understood in terms of an ordering principle, that of anarchy, which is the same for all international systems. Ruggie (1986) criticises Waltz’s use of ‘anarchy’ because there are no means of observing it. What can be observed are only the hypothesised effects of the anarchical ordering among states.

Ashley (1986) and Walker (1993), like other postmodern theorists, are critical of the attempts in IR theory to build ‘objective’ theory. They discount the possibility of any legitimate ‘objectivity’. Walker is critical of the emphasis that naturalistic IR theory places on ‘empirical or positive knowledge of what is’ (Walker 1993: 54, see also Zehfuss 2002). Walker says, ‘Claims about how the world is, or rather how it is to be known, become separated from claims about how the world should be. … facts are separated from values, empirical science is distinguished from normative theory and objectivity is opposed to subjectivity’ (Walker 1993: 54). On the critical realist side, Patomäki accepts the criticism as well because, ‘The data against which explanatory models are tested cannot be described independently of the theoretical language in use’ (Patomäki 2002: 90).

Responses to objections to the distinction

The first objection deals with the difficulty of drawing a sharp line between ‘observables’ and ‘unobservables’. Empiricists like Carnap hold that this is an inconvenience, but not a conceptual problem with the distinction. One standard argument uses the example of different colours. It is possible to categorise some objects as blue and some as green, even though there are objects whose colour is a shade of blue-green that makes them very difficult to categorise. Opponents will grant this. But the presence of blue-green objects does not show that the blue-versus-green distinction is invalid.

The second objection stems from the differences between scientists' and philosophers' usage. One of the difficulties that philosophers have faced in drawing the observation–theory distinction is that natural and social scientists use the terms 'observation' and 'theory' differently from philosophers, which causes some confusion. The philosophers' claim that there is a coherent distinction should not rest entirely on 'ordinary scientific language'. For IR theorists, to the extent they are aware of a distinction at all, they tend to use 'war', 'conflict' or 'crisis' as observation terms – much as physicists would say they observe the '1.06 second elapsed time it takes the ball to drop' as an observation. However, for the philosopher 'elapsed time' would be connected with an observation about the display numerals on a stopwatch, etc., while 'war' would have to do with humans in specified uniforms, travelling across specified territory, firing weapons, etc.

Usage, context, background assumptions and non-scientific grounds

Two problems arise from scientific usage, which are that it would be difficult to develop a coherent philosophical account of 'observation', based on scientists' usage of 'observation' and that it would be difficult to develop an account that involves consistency of usage. First, as far as coherence is concerned, philosophers widely agree that many of what scientists call 'observations' involve a considerable amount of inference. One observes instances in world politics of nuclear weapons states living at peace with one another. The United States and Russia are at peace today, as are France and the UK. But 'observing' these instances requires many inferences, since to say that peaceful Franco-British relations constitutes an instance requires inference from one's knowledge that both states possess nuclear weapons. This problem is severe only if one insists that a philosophical explication of the distinction must take into account how scientists in fact use the terms.

Second, as far as consistency is concerned, what scientists call 'observation' varies from context to context and is often relativised to the experiment at hand. The design of the experiment will involve background assumptions, i.e., the acceptance of propositions that are not being tested by the experiment. As Duhem (1954) has shown, one cannot test all of one's beliefs at the same time. So one must make assumptions about the functioning of a stopwatch or a telescope. One 'observes' that the object took 1.06 seconds to drop, but this requires the belief that the timing mechanism was working properly. Jerry Fodor's answer to this (1984), discussed below, is that while there certainly are theory-relative distinctions between observable and theoretical claims, that fact does not demonstrate the impossibility of there being a distinction between 'observation' and 'theory' that is *not* theory-dependent.

One might try to use language as the starting point in allowing ontological commitments. In that vein, what people talk about as real should be taken as real. While the state would emerge as a legitimate entity on this approach, it would be hard to distinguish in general what one wishes to call 'legitimate' entities from others. This 'linguistic realism' is justly criticised by Patomäki and Wight (2000). The purely linguistic attempt does not accurately get at the core of the distinction one should seek.

The third argument, based on the theory-ladenness principle has at least two versions. One is drawn from meaning-holism or the IP thesis and the other from the cognitive psychologists' claim of the continuity between perception and cognition. The IP argument (discussed on p. 40) aims to show that the interconnections between all of the elements in the belief-system of an individual or a scientific discipline are so intimate that any new belief entering the system has potential ramifications for all of the others. On this view the meaning of a statement that expresses 'observational belief' arises only by virtue of its connection to the rest of the network of beliefs. So theory T1 and theory T2 are incommensurable because 'the observation' used to test T1 as against T2 turns out to be two different observations; each gains its meaning by reference to the rest of the elements of T1 in one case and T2 in the other case.

In Chapter 6 (pp. 180–2), grounds for accepting IP and the arguments against it are adduced. It is worth noting that it is possible to accept IP and still reject the meaning-holism critique of the distinction between 'observation' and 'theory'. Meaning-holism casts doubt on the independent sources for the meaning of observation terms. In this case, it is difficult to provide a semantic distinction between 'observation' and 'theory'. However, just because there are reasons to doubt the existence of semantic grounds for drawing the distinction does not mean that there cannot be other grounds, especially epistemic grounds, for such a distinction (Fodor 1984). The web metaphor that helps illustrate this account of knowledge allows differences between the centre and periphery. The beliefs at the periphery are justified differently from those elsewhere. Causal semantic theories, e.g., of the sort that Wendt (1999: 57–60) endorses, cast serious doubt on semantic-holism. Semantic-holism states that all of the semantic properties of a statement derive from that statement's connection with other statements in the theory or corpus of beliefs. The causal theories hold that some of the semantic properties (though perhaps not all) are derived by their connection to the world.[4]

Fodor's response to objections

The second variation of the argument for the theory-ladenness of observation holds that there is no sharp distinction between 'observation' and 'theory' because cognitive psychologists now agree that observation or, more specifically, perception, is not separable from cognition. They hold that perception involves the cognitive process of inference. But Fodor attacks this argument and shows that there are beliefs that are acquired 'by sensory/perceptual processes [which are] theory neutral' (1984: 24). Fodor, a scientific realist (see 1984: 26, n. 1), nevertheless finds the distinction compelling and, to an even greater degree, he finds the criticisms of the distinction lacking. Fodor holds that there is a continuum between observables and non-observables and so there is no clear point of demarcation. Although there are intermediate cases in a grey area, this cannot undermine the conclusion that there is a difference between the two categories.

Fodor considers the background information one has, which is present in an observer's cognition as he or she views the world, and he considers how speakers acquire an understanding of the meanings of words in their native language. These considerations allow Fodor to draw a distinction between what people learn that becomes part of their 'accessible background' and what they learn that does not. Fodor holds that much of what people learn does not become part of the accessible background, for example, what the average person learns about physical theory. It is argued below that IR theory can similarly be excluded from the 'accessible background' category.

For example, one learns at a young age that the earth rotates on its axis and revolves around the sun. When learning this one still sees the sun rise and set. One's knowledge that the earth rotates on its axis does not affect one's 'observation' of the sun rising. The important conclusion Fodor draws from examples like this (and the formation of the morning dew) is that some things that one learns do not affect how one sees the world. Some of the well-known optical illusions that endlessly fascinate cognitive psychologists still have the interesting effect even after one learns that they are illusions. So, it is incorrect to say that everything one may learn affects how one 'sees' the world. Some learning forms part of a person's accessible background and some does not … and what one learns about physics, Fodor argues, does not. Fodor shows that only if all learning had this effect would one lose the common ground necessary for 'theory-neutral observation'.

Some learning in the natural sciences, on this view, does not affect how one observes the natural world. Astronomy and IR are admittedly very different. Still, a parallel claim would seem to apply equally well in IR. When one learns about IR theory, one is likely to learn of the long-standing rivalry between political realism, liberalism, idealism, Marxism and other theoretical approaches. Consequently, that awareness of lack of consensus would suggest that one would be even less likely to use what one has learned about IR theory to 'see' things in a fundamentally different way when examining observable events (like seeing a person seated at a desk with a specified official seal or insignia, signing a piece of paper with the word 'treaty' written on it). Nevertheless, one who is immersed in IR theory is just as likely as one who is not to observe a head of state or government signing a treaty.

While Fodor is trying to vindicate a form of common-sense realism and SR, his arguments serve to advance the empiricist's favoured distinction between 'observation' and 'theory'. Fodor does not claim that his arguments prove once and for all that the observation–theory distinction is unassailable, but rather claims, very persuasively, that the arguments that have been put forth to cast doubt on it do not survive scrutiny. (See also Goldman 1999: 238–42.)

Dretske's distinction between observation and theory

Thirty years ago, as Hanson's criticisms of theory-neutral observation and Kuhn's critique of positivist method had gained momentum, Fred Dretske (1969) offered a way of distinguishing between 'observation' and 'theory'. One of the keys was his distinction between two senses of 'observe'. He uses the particular case of 'seeing' to make important points about observation.[5]

The distinction Dretske draws is between what he calls 'epistemic' and 'non-epistemic' seeing. He acknowledges that there is a sense of 'observe' or 'see' that involves interpretation or inference. But, in part by emphasising the continuity of human and non-human animal sensory apparatuses, he argues that it is possible to see things without seeing them *as* what they are (or as something one might mistakenly believe that they are). One might simply see them, particularly when one's attention is focused elsewhere. When one reads a book while taking a walk one might, for example, be able to see objects in the background without identifying the objects as anything in particular and without even being able to remember seeing them. In this sense one 'sees' objects in a 'non-epistemic' way. A cat may see a computer without any theory, inferences or beliefs about it and an infant may see its mother for the first time without any beliefs about her.

This sense of 'see' is clarified by means of Dretske's concept of 'belief content'. Dretske says that propositions may have positive or negative belief content, or no belief content at all. There are some statements about the subject, Vladimir, that are related to Vladimir's having or failing to have a particular belief. And there are other statements about Vladimir that are not so related. Statements in the first group have some belief content, which may be positive or negative. Those in the second group have zero belief content. Any statement of the form 'Vladimir sees …' does not entail any statement of the form 'Vladimir believes that …' whenever the former has zero belief content. When there is no belief which is such that Vladimir's having or failing to have it is logically relevant to the truth of a given statement, then and only then does that statement have zero belief content. This group of statements is important for Dretske's argument about 'observation'. For example, the statement 'Vladimir walked past Estragon' has zero belief content. There are no beliefs that are logically related to the truth of the statement in the sense that Vladimir need have no particular beliefs in order for the statement to be true; Vladimir might be unaware of having walked past Estragon.

Dretske distinguishes the logical relationship from the psychological relationship by showing that seeing X is logically independent of having beliefs about X. This independence is seen by noting that one might see X and have false beliefs about what X is; similarly one might have true beliefs about what X is. The essence of this primitive notion of 'seeing' is simply that when Vladimir sees X, all that is entailed is that Vladimir visually differentiates X from its surroundings (1969: 20). Vladimir may have many beliefs about X, but none of those beliefs is entailed by Vladimir's having seen X. (See also the account of Chisholm 1957: 143–8.)

Ordinary usage, according to Dretske, provides support for his position. If Vladimir says to Estragon, 'You must have seen my green shirt, I was wearing it when you spoke to me in my office this morning,' Estragon may disagree only by asserting that the conversation took place over the telephone and not in person, or that Estragon was wearing a blindfold all morning, or something of the sort. Otherwise, Estragon can only say, 'Well, I didn't notice it, but you are right in that I must have seen it.' Estragon cannot acknowledge that the two spoke face to face and that Estragon's sensory organs were unimpeded and in normal working condition *and* that Estragon did not see the shirt Vladimir was wearing.

What Dretske has accomplished here is a severance of the logical link between 'observation' and 'epistemic states' and, similarly, a severance of the link between 'beliefs' and 'inferences'. Natural and social scientists may typically carry out observations making use of beliefs and inferences. But Dretske has provided a solid argument in favour of theory-independent observation.

Dretske admits that the non-epistemic notion of seeing or observing does not fit with the way scientists themselves speak of 'observation', which connects to the second objection laid out at the beginning of this section. To account for this broader usage, he explicates the notion of 'epistemic seeing', in both a primary and secondary form, where both involve 'belief content' but the latter excludes the requirement that the object of observation be seen in the primitive sense. Non-epistemic seeing (seeing-n) is different from epistemic seeing. A scientist can see a water molecule, but cannot see-n a water molecule. As far as secondary epistemic seeing, one may see in a secondary epistemic way that Patrick is taller than Lydia without seeing in a primary way either Patrick or Lydia, e.g., by inspecting a photograph of Patrick standing side by side with Lydia. Similarly, one may see in a secondary epistemic way that the gas tank is half full without looking at the gas in the tank but by looking instead at the gas gauge.

Dretske holds that observation differs from theory and he provides a broad notion of 'observation' which does conform with scientific usage. He says that a person's observations 'are those pieces of information, P1, P2, P3,… which are (i) relevant, or thought to be relevant, to his enquiry, and for which (ii) the statement that S saw Pi is true' (1969: 205–6). This definition, combined with the account of the statement 'S saw Pi', renders impossible the prospect that one can have false observations. Only epistemic seeing can provide an investigator with scientific observations. Non-epistemic seeing by itself can never do so.

The different notions of 'seeing' are related in a clear order. Dretske defines first non-epistemic seeing (1969: 20), followed by 'primary epistemic seeing' in terms of 'non-epistemic seeing' (1969: 141, 152) and then 'secondary epistemic seeing' in terms of 'primary epistemic seeing' (1969: 153). Secondary epistemic seeing yields genuine observations because it shares with primary epistemic seeing 'a crucial characteristic: they manifest the direct, non-mediated, attainment of knowledge by the possession of visually conclusive reasons to believe' (1969: 211).

Dretske holds that there are observations made without the need for theory or beliefs on the part of the observer, but also accounts for observations that do require theory, e.g., epistemic seeing. Thus, according to Dretske, critics of the observation–theory distinction hold that all observations are theory-laden. They support this generalisation by citing examples of purportedly pure observations that turn out, upon inspection, to be theory-laden. But Dretske argues, as noted, that just because some purportedly pure observations turn out to be theory-laden does not entail that all do. And indeed, it is not the case that all do, since there are examples of observations in the form of ‘epistemic seeing’ that do not require theory.

Dretske’s account may be extended to include events and processes. He says that battles (along with games, performances, etc.) are among the things that may be seen in this direct, immediate, non-inferential way. It is possible not to notice some things, particularly objects that are in the background, objects that do not move, etc. But events involve change and as such are noticed – the dog scampering out the door, the soldier advancing across the field. Thus Dretske points out the parallels between ‘seeing things’ and ‘seeing events’.

Possible reservations about Dretske’s argument

For those who are reluctant to accept Dretske’s extension of his argument against the theory-ladenness of observation from the physical to the social world, or even to accept his argument generally, there are further grounds for supporting an ‘observation–theory’ distinction capable of allowing IR theory choice.

As far as the extension to social observations is concerned, suppose that the distinction can be supported only for physical objects. IR theory choice requires the use of non-theory-laden observation statements, which provide evidence that permits comparisons with respect to specific criteria. Theory choice will involve the interpretation of statements that include theoretical terms. Inferences are necessary to move from the theoretical concepts that are part of the theory to the observation statements that allow theory evaluation. If Dretske’s argument is not extended from physical to social phenomena, then the inferences would have to move through an additional step, first from theoretical statements, to ‘social observation’ statements and then again to physical observation statements. The difference here is that inference is necessary in a quantitatively greater way, not in a qualitatively different way, because inference is already required for theory choice, even if the extension of Dretske’s argument is accepted. If extension is permitted, then the theory, with its theoretical statements, may be evaluated by means of tests phrased in terms of ‘observable’ social phenomena. If it is denied, then the translation must go though one additional step and the erstwhile ‘observable’ social phenomena must be translated into physical phenomena.

For one hesitant to accept the force of Dretske’s argument and who thus insists that there is a theoretical inability to separate all traces of theory-ladenness from observation, another response is possible based on the need to stay focused on the central question of this study, namely, what value IR theory may have to the practice of foreign policy-making? It is important to remain focused on that central task in the midst of abstract considerations.[6] What are the implications of the attack on the observation–theory distinction and the consequent theory-ladenness of observation doctrine for the problems of theory choice? The theory-ladenness of observation principle asserts that theory-neutral observation is impossible, i.e., there are no actual or possible theory-neutral observations. To disprove the theory-ladenness principle (which CS and CC advanced in this study requires) it is necessary only to show that there are some possible theory-neutral observations, not that all observations are theory-neutral. Dretske's argument shows that there is a legitimate sense of theory-neutrality, as in the case of 'non-epistemic seeing', that salvages the distinction. Beyond Dretske's abstract argument against theory-ladenness, there is a practical argument that shows that the minimal effects of the theory-ladenness of observation on the task of theory-choice in IR and foreign policy-making.

The danger the theory-ladenness of observation principle poses is that observations are not 'objective' and the lack of unbiased data will unfairly favour the corroboration of the theory whose concepts the observations import and disconfirm all competing theories. Consider the observation, 'Freedonia is at war with Sylvania.' A researcher might observe the duly authorised body in Freedonia declaring war, might be on a battlefield and 'see' a battle, or might read a headline of war in a reliable publication. Any of these experiences would support the observation that Freedonia is at war with Sylvania. Consider next an eight-year old child who reads the headline 'Freedonia Goes to War With Sylvania' in a reliable publication. That child does not have any theory of war and peace. He or she may have background beliefs, which might include an ontology of social reality that encompasses nation-states in general, Freedonia and Sylvania in particular, and the possibility of violent relations between them (if relations are said to have ontic status). But the eight-year-old child (one hopes) has not read Thucydides, Hobbes, Kant, Marx or Waltz) and he or she may not have any theory of war.

Three points thus follow from the example. First, the child understands the key terms much as an IR theorist does. Although the child may not assign 1000 battle deaths as the minimum to qualify as inter-state war, neither does the author of the headline, in all likelihood, apply that criterion. Second, the child possesses no beliefs that constitute a theory of IR. The child does have some set of background beliefs in virtue of which the headline is interpreted but those do not constitute a theory of IR. Third, those particular beliefs do not preclude any theory of IR. The child may learn about Marxist theories of war the next day, but not be forced to surrender his or her understanding of the literal meaning of the headline.

One might argue that, theoretically, there is a sort of 'ladenness' at work, it is, nevertheless, false to say that it is 'theory-laden'. The interpretive scheme is so broad as not to qualify as a theory and so as not to be inconsistent with contending theories. If the sort of 'ladenness' is this limited, then it is reasonable to think that the problem of the theory-ladenness of observation principle biasing theory-choice in IR would not arise. Consider the theories of the authors just cited: Thucydides, Hobbes, or others. How would the ontological commitment to nation-states and violent inter-state conflict bias theory choice? All of these theories include such entia (though there are some that do not). Therefore, the problems for practical need of theory choice seem minimal.

It is important to remember that no matter how scientific one's enquiry, whether social or natural, theory choice is always conducted at some time, some place, with a set of observations at hand and a specific set of competing theories. What follows from worries about the theory-ladenness of observation principle is that, as theoretical work proceeds, the language of the observations must be carefully monitored so as to avoid prejudice between any of the contending theories. It is possible that theory-laden language may be employed, so long as the observation language carefully remains neutral as between the particular theories under review.

The presence or possibility of some theory-laden observations does not establish the inherent theory-ladenness of observation. The general principles assert, contrary to what is shown by the above argument, that all observations, without exception, are theory-laden. For those who do not find the argument persuasive, the more important point is that for the practical purposes of policy-making, the theory-ladenness of observation principle does not undermine unbiased theory choice. So the critic who uses the theory-ladenness of observation as a basis for an attack on the possibility of unbiased theory choice would have to show not only that theory-neutral observation is in principle impossible, something that Dretske's argument seems to preclude, but also that if the theory-ladenness of observation principle is true, it has real effects on theory choice in IR.[7]

The literal truth and falsity of scientific laws

Chapter 2 noted that some authors see a divergence between the natural and the social sciences in that only in the former do the best theories have genuinely true laws. They hold that social science theories have laws that are at best idealisations or approximations because the social universe is so complex and has so many variables that come into play across a range of cases. These analysts hold that social science laws may offer some guidance, but they are not precise and are not literally true. Hence naturalism fails because of this crucial difference in the nature of laws in the natural versus the social sciences. Chapter 2 cited examples like the generalisations of Waltz regarding balancing of states or of Keohane regarding the cooperation-inducing effects of increased distributions of information of regimes. This book defends a fairly strong version of naturalism. Hence it is important to show that, as Waltz says, 'scientific theories deal in idealizations' (1997: 914).

Reflectivist critics of naturalist IR attack the 'lack of accuracy' of the laws and theories in IR. For example, Patomäki (2002: 75) criticises Singer for his emphasis on explanation and description that does not hold accuracy as a goal that cannot be compromised. Patomäki cites Singer (1961: 79) as saying 'the primary purpose of theory is to explain and the descriptive and explanatory goals of the researcher may come into conflict. When that happens, the researcher must give preference to the explanatory goal, even at the cost of some representational accuracy.' Chapter 1 postulated that it is unreasonable to hold the social sciences to a standard of scientific theorising that the natural and especially the physical sciences do not meet.

Furthermore, the physical sciences do not offer laws that are representationally accurate, or known to be true. As such, the laws of the natural sciences are similar to, and not in contrast to, laws of the natural sciences with regard to accuracy. The case is best made by Nancy Cartwright (1983) who, as noted above, clearly distinguishes two kinds of laws. Observational laws describe what is experienced and observed and theoretical laws 'explain' observational laws (though Cartwright calls them 'phenomenological' and 'fundamental'). She argues that the cost of explanatory power, which is generated by the theoretical (fundamental) laws is descriptive accuracy. One of Cartwright's examples from physical science is quantum damping. She notes that there are texts (like Agarwal 1974) in which many different sets of laws (six in Agarwal's case) are offered, based on various different sets of mathematical equations (three by Agarwal), to explain the same phenomena. Physicists, according to textbook accounts, offer these mathematical explanations for the phenomena but do not attempt to show which treatment is 'true'. However, she cites the author of such a textbook who says that, for practical purposes, the range of laws and equations complement one another: '[d]ifferent approaches are useful for different purposes' (Cartwright 1983: 81).

Cartwright says that the use of approximations does not come about because the 'true' mathematical statements are too difficult. The process scientists use seems to indicate quite the opposite. She says, 'The steps in a derivation move away from the rigorous consequences of the starting laws, correcting and improving them, in order to finally arrive at an accurate description of the phenomena' (1983: 14–15).

A proper understanding of how explanation works, according to Cartwright, shows that laws are not directly connected to reality. Such a connection is part of a common misunderstanding of 'explanation' and from it arises the appearance of truth. She proposes a simulacrum concept of a model, according to which the simulacrum has form or appearance without having its essential structure. So to explain is to fit the phenomena into a theory. But the fundamental laws of the theory apply only to the objects in the model, not to the real-world phenomena. The objects in the model have only the form or appearance of the real objects without having the proper qualities.

For Cartwright the explanatory law in a theoretical explanation may very well do its work and succeed in meeting proper criteria for acceptability among laws, despite the fact that the law is not true. (Causal laws are exceptions to this claim.) She says, 'What is important to realize is that if the theory is to have considerable explanatory power, most of its fundamental laws will not state truths, and that this will in general include the bulk of our most highly prized laws and equations' (Cartwright 1983: 78).[8]

Cartwright's distinction among theoretical laws between mathematical and causal includes the idea that only in the case of the former may there be multiple models, which are useful for different purposes and which bring out different aspects of the phenomena. With regard to the former class, 'Which is the right model?' is the wrong question, since different models bring out different aspects of the phenomena. This contrasts with how investigators proceed in the case of causal explanation, where one does not begin with one causal story and then, out of convenience, switch to another. She says the investigator must produce one causal story and hold to it. These stories are treated as true and false. (There is tension between causal and theoretical explanation. The causal explanations show how various causes combine with one another, while theoretical laws are needed to show just what each cause contributes.)[9] In the social sciences, rational choice explanations, which posit the existence of a rationally behaving state and the rational economic person, are clearly idealisations. They are like – and not in contrast to – laws in the physical sciences, laws that make use of the notions of ideal gases or frictionless machines.

Selecting a theory

Acceptance of the theory-ladenness of observation principle and IP make it more difficult to base theory choice on rational grounds. The standard interpretation of IP is that theory choice and the history of science must be accounted for outside of discussions of strictly rational standards. The above defence of the observation–theory distinction, the meaningfulness of at least some 'theory-neutral' observations and the rejection of the theory-ladenness of observation principle, change the terms on which theory choice may be defended.

Criteria of theory choice

As policy-makers select among options they compare their various predictions of what is likely to happen should they pursue each of the possible courses of action (or inaction). They ponder how the future will be different under each of the options under consideration. These predictions are justified by the set of causal and constitutive relationships they accept, which, if sufficiently systematic, may constitute a theory. But there are many theories of IR and they generally yield different predictions about how one action (ushering in democracy in Iraq) will produce results and how the results (like peaceful relations with other democracies) would stack up against those of another policy action (allowing Iraqis to choose a form of government which may not be democratic). Which policy one believes will bring about the desired goals will depend upon which theory of IR one accepts. Given a fixed set of objectives for the state, like reduction in terrorist operations against the US, in a given set of circumstances, like the conditions that obtain currently in the Middle East, a liberal internationalist theorist who supports Russett and Oneal's (2001) view of DP relationship might compare three policy options for the future of Iraq and conclude that ensuring a transition to democracy in Iraq will promote peace with the US, Western European states, Israel and other democracies. A classical realist, on the other hand, may conclude that ensuring a pro-American leadership, even if not especially democratic, is the most effective way to enhance US security.

How then does one choose a theory? That is, what are the criteria of theory choice? Clearly, scientific or rational enquiry cannot do without some set of criteria. Theories are abandoned when new theories take their places; and if this is a rational process (and there are those who deny it is, such as Kuhn 1962 and Feyerabend 1978), then there are rational criteria for making the choice. Various criteria have been cited over the years by philosophers of science.

Political scientists like Vasquez (1998) and King et al. (1994) have defended falsifiability, consistency, concreteness and breadth or range. Some of the other criteria cited by philosophers of science are: fecundity, degree of corroboration, methodological conservatism, i.e., minimal change relative to previously accepted theory, etc. While many criteria have been advanced by various authors, the most universally accepted criteria are: internal consistency, coherence, simplicity and explanatory power. Studies of the history of science seem to reinforce the notion that these four are invoked by scientists.[10]

When one considers the debate discussed in Chapter 2, it is interesting to note that only anti-scientific realists have no difficulty offering rational justification for all four of these criteria, while scientific realists have considerable difficulty justifying the widely endorsed simplicity criterion. There was no difficulty justifying it when philosophical realism was based on theology: God is omniscient, omnipotent and does nothing in vain and hence would not create a world of complex laws and extra or unnecessary forces when a simpler one could perform the same functions (see Leibniz 1956). In the secular world of science, the grounds for the simplicity criterion have been much harder to come by for scientific realists. Why must the reality that our theory reflects be the simplest one consistent with the observations? If it is science that one hopes to be able to account for, theoretical entities are much easier to do without than the criterion of simplicity.[11]

The criterion of 'simplicity' has had a significant influence not only on philosophers' attempts to construct a normative account of science, but on the history of science directly. Conservation of motion was seen as a simpler understanding of the physical world than its negation. Historically, the theories of Newton and Descartes adopted different definitions of 'motion'. Newton's theory was finally accepted over Descartes' in large part because only Newton's concept of 'motion' fulfils the 'conservation' desideratum. To the extent that this was a factor, the criterion of 'simplicity' was used to help define an observational term ('motion') for theoretical purposes.[12] This example suggests how important simplicity is to science. In a sense it has proven to be more basic than the distinction between 'observation' and 'theory'.

While scientific realists are happy with the rejection of the 'observation/theory' distinction, the importance of the 'simplicity' criterion poses a real problem for the scientific realist account of science, one that the conventionalism advocated here solves.

For those who take the efficiency of deriving conclusions as central to their theory of science, there is little problem in justifying the criterion of 'simplicity', which is clearly a contributor to convenience. The attitude of Mary Hesse is typical, when she says which system of laws should be accepted 'will always depend on what system of laws is most convenient, most coherent and most comprehensive' (Hesse 1980: 71). Similarly, the criterion of 'methodological conservatism' or 'minimal change', while not one of the most ubiquitous criteria of scientific theory choice, poses the same problem for the scientific realist. It is hard to produce reasons why continuity with older theories should give the theorist a more direct grip on the truth. Why is the theory most like the one previously accepted more likely to be true than one that is otherwise equally powerful and appealing?

'Diversity' is different from simply 'number of confirming (or non-falsifying) instances'. Thus Vasquez is not in accord with the principal view in the philosophy of science when he says: 'A theory or research paradigm that has the most corroborated hypotheses and least anomalies is obviously the best or most promising one to use in order to achieve the purpose of science' (Vasquez 1998: 31). The above argument shows that this claim is overstated, since the 'most corroborated hypotheses and least anomalies' have to be counterbalanced against other criteria, especially 'diversity of phenomena'. A theory with many corroborated hypotheses, where those hypotheses are similar and account for very similar sorts of cases, is less impressive to most methodologists than another theory whose fewer corroborated hypotheses cover a wider array of types of situations (or ranges of values of the variables).

Attempts to show that scientific theories have inherent or purely logical features on the basis of which the rational investigator can univocally choose one among the competitors have been regarded as failures. Those who follow Feyerabend, Kuhn or Quine in adopting the IP thesis or the radical underdetermination principle contend that theories have no inherent features that allow investigators to pick out the unique best among them and that there is no relationship between theory and evidence that would allow them to select the best theory. This issue is dealt with more fully in Chapter 6. For present purposes, it is helpful to show that there are considerations of a rational criteria of theory choice that can be found outside of the theory and its direct relationship to the evidence.

Truth, theory choice and pragmatic considerations

Pragmatic considerations have proven to be very powerful in solving important problems and in stimulating a great deal of theorising in the philosophy of science. The pragmatic approach has the advantage of making sense of scientific practice, of accounting for the desiderata of 'simplicity', 'explanatory power' and the like, and even of retaining inference to the best explanation (IBE), for those who find it compelling. This may seem unlikely, since IBE is usually identified with SR, as it was by Wendt (see p. 54) and SR is usually defined as committed to a correspondence and not a pragmatic theory of truth. The pragmatic, non-scientific realist approach adopted in this book generally eschews a correspondence theory of 'truth' (Quine and Ullian 1978).

The pragmatic theory of truth, which is endorsed here, is tied to the doctrine of motivational realism (see p. 95). While it is possible to believe that there are real subatomic particles and international political structures, it is possible also to realise that human knowledge is ultimately bound up in human practice and goal-directed activity. Thus the ability to achieve goals through action is the ultimate test of 'knowledge'.

IBE may be seen as acceptable to non-scientific realists in the more circumscribed form Peirce endorses because it adds to methodological convenience. As noted above, there is much intuitive plausibility to IBE, but the dangers of error of the attendant commitment to SR, particularly social SR, were too great to warrant it. However, there is a more modulated form of IBE, in which it functions as an intermediate rather than terminal step in enquiry, that has greater plausibility than the simpler variety to which contemporary analysts usually refer.

Peirce was the first to formalise the notion of IBE, which he called 'abduction'. Peirce's highly influential account of science holds that IBE begins scientific enquiry. On the basis of IBE, an hypothesis is selected, which accounts for or explains existing observations. Deduction is next used to derive empirically testable consequences from the hypotheses. If the tests prove to corroborate the hypothesis, induction is then used to argue in favour of the hypotheses.

Peirce holds that hypotheses and observations are clearly connected to one another. Many have argued that such a connection requires acceptance of the analytic–synthetic distinction. Others, like logical positivists, have argued that such a connection requires that one regard as strictly scientific forms of enquiry nothing but physical sciences. There is a serious problem in maintaining the connection, because i) most analysts concur that giving up all non-physical science enquiry is an unacceptable consequence and ii) the criticisms of the analytic–synthetic distinction by Quine (1953) and subsequent figures have been regarded as quite successful. The position that Peirce develops (1932: para. 780 'notes on ampliative reasoning') allows one to maintain the connection without either having to accept either i) or ii).

Inference to the best explanation as an element in anti-SR accounts and Peirce's solution

Most scientific realists endorse IBE and the latter does seem to possess considerable plausibility. When a person comes out to a car sitting for a year in a locked garage and observes a partially flat tyre, he or she may explain its condition by considering that it has a slow leak. Since the car has not been driven it is unlikely that it has suffered a large puncture. So the slow-leak hypothesis is the best explanation for its flat tyre. One accepts the explanation that there is a slow leak and proceeds to fill the tyre with air rather than take some other course of action, such as ordering a new tyre from a store or filing a vandalism complaint with the police. Why not do the same thing in science? On Peirce's account, scientific reasoning goes from abduction, to deduction, and then to induction for further corroboration of the hypothesis. This view seems to accord with the flat tyre example. The careful motorist will, after filling the tyre with air, keep a close watch on that particular tyre in the hours and days that follow to see if it deflates (i.e., he or she will draw deductive consequences from the best explanatory hypotheses and look for inductive support).

Peirce argues vigorously in favour of the importance of abduction or IBE as a part of the method one should use to attain the highest form of empirical knowledge, namely, scientific knowledge. While always fallible, science is capable of providing very high-quality knowledge. IBE is an essential part of that methodology. But for Peirce, IBE is not a process that allows one to justify belief in an hypothesis or theory; IBE is a form of inference that justifies the testing of that hypothesis or theory. If the hypothesis or theory does not offer an explanation of something one seeks to explain, it is not worth testing. But on IBE, the success of the hypothesis, or its ability to explain, is not a ground for accepting its truth. Its acceptance can come only when it has been successful in meeting the criterion of 'best explanation', followed by the deductive derivation of consequences or predictions from it, which are then subjected to empirical, inductive tests. For Peirce, who can be considered a naturalist and a supporter of common-sense realism and SR, scientific reasoning for the purpose of theory choice will always involve abduction, deduction and induction.

Peirce believes that good theories are true and help scientists to learn about an external reality. But the pragmatism (or 'pragmaticism', as he later called it) he developed entails an account of 'truth' more complex than correspondence or coherence theories. As noted, scientific realists generally endorse a correspondence theory, which holds that a sentence is true if and only if it corresponds to an external reality or independent fact of the world. Coherence theories of truth hold that a sentence is true if and only if it is a part of a system of sentences that is coherent and consistent. This view allows for the possibility that there may be many different sets of sentences that one might accept, since the sentence need not correspond to any independently existing reality.

The pragmatic theory of 'truth', which is now widely adopted and which Peirce developed, holds that truth lies in the conformity with experience. To say that a statement is true is to say no more than that the statement conforms with experience. Anti-scientific realists who have followed Peirce on this point have noted that this conception of 'truth' does not entail that there is an independently existing reality (i.e., does not entail common-sense realism) or that all the terms of scientific theories refer to extant entities. Peirce did, however, hold that there is an independent world and that a metaphysical structure undergirds the world that appears to us in experience. However, when he holds that scientific statements are true, he means precisely that they conform with scientific experience.[13]

Because scientists do not act on their theories Peirce holds that they do not have to believe or accept their theories. While Peirce fits the general definition of 'scientific realist', he does not agree with most contemporary scientific realists on this point. According to Peirce's pragmatic theory, belief is a 'habit of action'. To say a person believes P is to say that he or she is willing to act on the basis of P. If such willingness to act is not forthcoming, then whatever he or she may say about his or her belief with respect to P, he or she cannot be regarded to hold that belief. An engineer, unlike a scientist, must act practically, and so must believe or reject scientific propositions and theories (Peirce 1932: 5.589). Hookway (1985: 73) distinguishes Peirce's fallibilism from scepticism only by Peirce's tenet that one could be assured of arriving at the truth if one had no limit on the time and effort available for investigation. Rational self-control has no place in settling issues of practical importance (Hookway 1985: 74).

So one must conclude that even the philosophical and scientific realist need not hold that 'truth' must be viewed as correspondence or that theories need be 'believed' by theoreticians. It is the practitioner who must believe and draw conclusions. But even for such a person, Peirce did not view IBE as warranting acceptance or genuine belief of the best explanatory propositions, or as warranting belief in the existence in the entia posited by the best explanatory theories. It was just one step in the process required for settling on such beliefs.

Multiple almost-equi-probable hypotheses

One problem with IBE (shown in Chernoff 2002) arises from nearly equi-probable hypotheses. The problem appears whether IBE is viewed as inference to the best theoretical or best causal explanation. Consider the case in which the investigator knows that the observed phenomenon e is more likely to have been caused by a than by b or c. But he or she also knows that 35 per cent of e's are caused by a's, 33 per cent of e's are caused by b's and 32 per cent of e's are caused by c's. In this case it seems that an agent with a practical job to do – an engineer or policy-maker – may learn of the possibilities and may have to make a decision to act. Action then might favour slightly the belief that e was caused by a. (This arises because the actor must act and choose one, even when the probability of all three is 33.3 per cent.) However, the theoretician, who considers epistemic grounds, has no practical need to act. The theoretical investigator lacks good (epistemic) grounds to conclude that a caused e. The investigator knows that either a or b or c caused e, that e's are present and so a's or b's or c's must be present. But he or she has no good epistemic grounds to infer that a's are present in the case described.

This objection to IBE does not work against Peirce's position because Peirce does not see abduction or inference to the best explanation as a producing conclusions to be accepted. Peirce sees IBE as generating hypotheses that warrant further testing, which seems entirely consistent with one's intuitions when all the investigator knows is that causal breakdown is 35 per cent, 33 per cent and 32 per cent. Furthermore, the problem is easily handled when the distinction between theoretician and practitioner is drawn.

Before leaving the consideration of 'truth', it should be noted that while SR is generally associated with a correspondence theory of truth, the closely related doctrine of critical realism is not. Patomäki, drawing on Bhaskar, defines truth as follows: 'Truth is a regulative metaphor, which has normative force. Truth is a human judgement, which is based on a *metaphor* of correspondence to the way the world really is' (Patomäki 2002: 14, see Chapter 5). From this quotation it is unclear whether truth is a metaphor or is based on a metaphor. These are two very different things. Later Patomäki, following Harré , says, 'truth relegates the notion of correspondence between epistemic statements and ontic objects to the status of a metaphor for the goal of scientific practices, which aim at articulating a matching representation of components of a complex and their essential relations' (2002: 148). Patomäki adds that truth has 'normative political implications' (2002: 143) and that 'truth is one value among many' (2002: 160).

Patomäki quotes Bhaskar's comment that on Popper's account it is 'decisive to realise that knowing what truth means, or under what conditions a statement is called true, is not the same as, and must be clearly distinguished from, possessing a means of deciding – a "criterion" for deciding – whether a given statement is true or false' (Bhaskar 1986: 100–1, cited by Patomäki 2002: 161 n. 2). Yet Patomäki goes on to criticise Popper as follows: 'The problem is that it presupposes that the truth is already known, for unless the truth is already known, how would it be possible to tell something about the truth-content of an assertion, statement, belief and the like?' But Popper is trying to distinguish what 'truth' is from what can be known to be true. Indeed, Patomäki says, 'Any notion of truth implies criteria, standards and measures, which have social and political consequences' (2002: 143). He thus appears to commit what he and other critical realists, following Marx, call the 'epistemic fallacy', according to which 'what is known is what can be … observed and what "is" is what can be known' (Patomäki and Wight 2000: 217).

Conclusion

The theory-ladenness of observation principle states that all observations are theory-laden and that there is no theory-neutral observation, which undermines the traditional, rationally based naturalist approach to theory choice. The theory-ladenness of observation principle forms the basis of an attack on theory choice because, if it is accepted, the terms of behavioural laws like LB1 would be understood sufficiently differently by advocates of the competing explanatory laws LE1a and LE1b so that no rational criteria for resolving the dispute between them would be possible. That is, if the theory-ladenness of observation is admitted, then there is no way to resolve resolution of disputes between advocates of LE1a and LE1b (or between advocates LE3a and LE3b, etc.) because the disputants are not even attempting to explain the 'same' law. This chapter has shown, primarily by examining Dretske's argument, that there are reasons for rejecting the theory-ladenness of observation principle and thus for distinguishing observation statements from theoretical statements and observational laws from fundamental laws. The observation–theory distinction allows investigators to agree on observed patterns of behaviour (summarised in observational laws like LB1, LB2 and LB3) and to debate how those regularities should be explained (whether along the lines of LE1a versus LE1b, LE3a versus LE3b, etc.).

This study defends the value of an empirical and causal approach to IR, along with the importance of explicating constitutive relationships. One's ontology should include genuine social entia. But as was argued (pp. 52–8), there are not adequate epistemic grounds for believing that the 'real' theoretical social entia and forces have been properly identified (see also Chernoff 2002). While these entia are 'constructed', it does not follow that investigators have privileged or unimpeded access to them. They do not provide the investigator with indubitable statements. Mistakes about these entia remain possible, and knowledge of them remains fallible. The difficulty of choosing between the different, competing theoretical claims about such entia underscores the dangers of the scientific realist's and the critical realist's practice of inferring the real-world existence of theoretical entia to which one's preferred theory is committed.

This study advances a causal-conventional account of IR meta-theory, which requires the use of a distinction between observation and theory. This chapter has sought to examine laws in the social sciences and show that arguments against the distinction do not withstand scrutiny. Consideration of the observation–theory distinction in the physical world was extended to the social world. The discussion of 'convention' and 'consensus' in Chapter 6 makes it clear that there is a possibility of natural-science-style approach-to-consensus (see also Chernoff 2004). It rejects limitations on the corpus of justified beliefs stemming from Kuhn's IP thesis and Quine's radical underdetermination principle. Nevertheless, Kuhn's distinction between mature and immature disciplines must be taken seriously, especially when appraising the state of IR theorising. IR theory is decidedly immature compared with that of physics.

Although IR naturalists affirm parallels with the natural sciences, they must admit that they are not identical. Each theorist who accepts some form of naturalism has to identify limits to naturalism and show that there are differences between theories in IR and those in physics. Little (1993a, 1993b) talks about the multiplicity of forces in the social world, that is, the complexity of causes in the 'open system' of social action; Bohman (1993) focuses on the inescapability of the hermeneutic circle; and Wendt (1999) notes the asymmetry of supervenience, where social factors supervene on the physical but not vice versa. The fallibilist account offered here stresses that the investigator has limited access to social reality. To be sure, the world that social science studies is complex. But the physical world is, too. (For an interesting argument that the problems of physics are simpler, see Bernstein et al. 2000). Newton had to invent new and very complex mathematics to be able to formulate his theory, and he had to use very complex applications of that mathematics (like second derivatives) to describe the physical world. This is far from simple, and the philosopher of the social sciences must not underestimate the complexity of the physical world. The naturalism advocated here acknowledges the complexity of both the natural and social worlds.

The present account of IR relies on the theory of knowledge in a central way. Scientific realists in IR (like Wendt and Dessler) and critical realists (like Patomäki and Wight) want to replace this focus with a focus on the theory of being. But this study presents reasons to conclude that, from the point of view of the foundations of social science, the evidential access to different sorts of putative entia, especially physical versus social entia, distinguishes the investigator's grounds for believing propositions about physical entia from propositions about social entia.

The present chapter shows how scientific realists, critical realists and other critics of traditional naturalism have based their attacks on the alleged universal impossibility of distinguishing 'observation' from 'theory' and on the argument that the social sciences, unlike the natural sciences, produce inaccurate laws. This chapter defends naturalist methods by showing the theoretical and practical failure of the attack on the 'observation–theory' distinction and the approximations inherent in both natural-science and social science laws. The conclusions drawn early in this chapter pave the way for the defence of a set of rationally grounded criteria of theory choice, offered in the preceding section. The next task in establishing a basis for a discipline of IR with practical value is to examine perhaps the most perplexing aspect of natural-science or social science theories, namely the 'causal' character of the laws they contain.

4 Natural causation, social action and international politics

'I would rather discover one cause than gain the kingdom of Persia'.

Democritus, translated by Diels, Fr. 8 (cited in Freeman 1948: 104)

SGANARELLE: My advice is that she [having apparently been struck dumb] be put to bed again, and, for a remedy, you must make her take plenty of bread soaked in wine.

GÉRONTE: Why so Monsieur?

SGANARELLE: Because in bread and wine when mixed together there is a sympathetic virtue which produces speech. Do you not remember that they give nothing else to parrots, and that it teaches them to speak?

GÉRONTE: Oh! What a great man you are! Quick, bring plenty of bread and wine.

Molière, *A Physician in Spite of Himself*, Act II, scene iv

RAY: Who knows why anything happens, Theresa?

David Rabe, *The Dog Problem,* Act II

Why did Saddam Hussein expel UN weapons inspectors in 1998? Why did the Nazis invade the USSR in 1941? The study of IR is replete with 'why' questions, and such questions are usually understood as questions of causality. Moreover, as noted in Chapter 1, policy-making requires justified beliefs about the future and about how different policy options will bring about change. People ordinarily understand these connections as matters of causal efficacy. This chapter considers what is involved in making and justifying such claims. Because both explanation (which seeks to answer 'how' and 'why' questions) and policy-making (which assumes predictive capacity) are usually justified by some notion of 'causality', it is essential to gain an understanding of the role of that notion in IR.

Authors who examine the invasion of, and nation-building efforts in, Iraq present many causal claims, either implicitly or explicitly. Daniel Byman (2003) offers an account of what causal forces will need to be applied to create a democracy in Iraq. He says that a transition to a stable, democratic state in Iraq will require 'massive help from the US and other powers' (Byman 2003: 49). Byman's analysis accords with that of the Rand team assembled to examine prospects for nation-building in Iraq (Dobbins et al. 2003). The Rand study cites interference from neighbours as an additional independent variable, while Byman argues that interference from neighbours, which he also regards as extremely damaging to the democratisation efforts, is a result of a lack of US and Western aid. The Rand study cites a number of causal generalisations as particularly important. Examples include: cabinet departments work harder to promote projects that they see as part of their core mission than for others (Dobbins et al. 2003: 221). Insecurity causes continued political turmoil; that is, security is a necessary condition for political development (2003: 69). Neighbouring states can cause nation-building efforts to fail, if they choose to do so; that is, successful nation-building requires regional cooperation or, at least, acquiescence, since a fragmented nation cannot be rebuilt if neighbours are trying to de-stabilise it (Dobbins et al. 2003: 107, 166). The level of effort is the greatest controllable factor determining the ease of nation-building (Dobbins et al. 2003: xxv). The higher the proportion of stabilising troops, the lower the number of casualties suffered and inflicted (Dobbins et al. 2003: xxv).[1]

Hollis argues that heavy-handed policies or extremely visible presence of the US will unite all of the opposing forces (Hollis 2003: 32–3), which appears to function as a causal law. According to Hollis (2003: 25–6) theories are not entirely universal, in that an adequate theory must take into account the historical peculiarities of a region. Theories that lack these dimensions will fail, as evidenced by their predictions. A similar analysis is offered by Brooks (2002), who sees a ‘groundswell’ of support in the Middle East for Islamist ideals. She offers an analysis with a set of causal statements, from which she derives predictions and prescriptions. In Brooks’s view, if Arab regimes liberalise, then Islamist groups will lose ‘their most effective rhetorical devices’ (2002: 616). Because the groundswell metaphor presents Brooks with the image of an unstoppable force, the question that remains is how to direct or channel that force. With the groundswell metaphor, the force is there – and has to be dealt with rather than perhaps made to disappear. However, a different metaphor might offer more hope to those who seek to repress Islamist movements.

For example, Metz (2003–4: 25) offers the metaphor of a cancer and, accordingly, holds that if the danger is identified and dealt with ‘at a point early enough in [its] development’ the danger can be minimised. Metz cites a number of causal factors that the US must consider as it plans for a transition in Iraq. He focuses on the issues connected with quelling the insurgency in Iraq. Metz (2003–4: 28) says that the ‘raw materials’ are anger, resentment, alienation, frustration and a unifying ideology (the latter of which is absent) among Iraqi opponents of the US and that they are causally linked to a forceful insurgency. He adds that financial resources for the insurgents also increase their ability to recruit new members (Metz 2003–4: 29). Metz sees a complex causal picture with essentially two very different models of insurgency. One seeks to control territory and confronts an ethnically similar enemy, such as the Communist Chinese in the 1930s and 1940s. The other seeks to expel a foreign occupier but does not seek incremental gains in territory, such as the Palestinians today.

Barton and Crocker (2003) argue that ten major steps must be taken to bring about stable democracy in Iraq. They base their analysis on several previous attempts at post-intervention stabilisation, especially Bosnia, Kosovo and Afghanistan. Excessive financial debt burden will cause economic hardship in Iraq and open the door to political instability (Barton and Crocker 2003: 9, 20). They say (Barton and Crocker 2003: 12) that delays in securing internal order will allow destructive power-struggles among Iraqi factions and that slow progress towards external security will encourage nervous neighbours, especially Turkey and Iran, to take steps inimical to Iraqi security (Barton and Crocker 2003: 14, 21). Lewis (1990, 2002) offers an analysis of Western relations with Arab states that focuses on the increasing success of the West on social, cultural, scientific-technological and economic grounds over the past five centuries. He emphasises Middle Eastern Muslims' growing sense of injustice at the hands of the West, which leaves the US a choice between disengagement and involvement in creating democracy (Waldman 2004). Lewis's analysis offers the following generalisation: a group with a unified identity that represents a large portion of the world's population, will, if it has an historical sense of injustice, seek to equalise or redress that inequality possibly with violence (Waldman 2004).

The view of causality advocated here rejects causal scepticism and acknowledges the meaningfulness of causal statements. The preceding chapter defended Dretske's claim that individuals can observe and experience wars, famines and chess matches. Events can be the object of an observer's experience and can be associated to like events, e.g., to other wars, to other famines and to other chess matches. A class of such events constitutes an 'event-type'. This chapter defends a version of Duhemian conventionalism which is consistent with an account that takes seriously causation among event-types and that leaves open (and thus need not deny) the principles of SR.

There have historically been problems associated with 'causality', typically treated in the context of physical objects and processes and there are specific problems of attributing 'causality' to social phenomena. For this reason the aim of this chapter is to consider the problems of event-types and of context-dependence (pp. 90–3); to note how the problem has changed in the past century because of modern physical theory and non-deterministic or probabilistic generalisations (pp. 93–100); to defend a version of 'conventionalism' in scientific enquiry that is applicable to both natural and social sciences (pp. 100–17); and to address the objection to a naturalistic view, crucial for IR and social science analysis, that reasons cannot properly be understood as causes (pp. 117–23).

The problem of devising an account of the notion of 'causation' has proven to be one of the most vexing problems in the history of philosophy. Hume says in the *Treatise* that no question 'has caus'd more disputes among both ancient and modern philosophers than the relations of cause and effect' (1965: 156). This extremely problematical notion, some empiricists (e.g., Russell 1918) have argued, should be banished. It connects to most of the major problems in the history of philosophy, such as the mind–body problem, the nature of substance, the omnipotence of the Creator and the limits of human knowledge. The account offered here is an attempt to deal with a subset of the standard questions surrounding the notion of 'causality', particularly social causes and their relationship to natural causes. The focus is on how these pertain to the central problem of this study – the policy value of IR theory. In any case, one must not lose sight of the immense difficulties and complexities involved in a comprehensive answer to the problem of 'causality'.

Event-kinds and causal context

Events and event-kinds

The first question to ask is what the things are that causal relation statements are talking about. What is it that causal relations relate? To put it another way, causal relations are relations between what sort of things? The ancient views tended to assert causal relations between substances or bodies; one thing was the cause of another thing. But this view was largely supplanted by the view that one event causes another. Some, like Carnap (1966: 190), also argue that processes, facts or propositions are the relata of causal relations. Salmon (1980) argues that 'statistical relations constitute causal relations' (see pp. 93–7 below). In any case, the most widespread view today is that 'the assassination of Archduke Franz Ferdinand' is not a thing, substance or body but an event. Franz Ferdinand and a match are both things. But Franz Ferdinand did not cause the war nor did the match cause the fire. Both the Archduke and the match existed quietly for some time without the outbreak of wars or fires. And one can imagine Franz Ferdinand and the match continuing to exist and finally meeting their respective demises of natural causes without any world wars or fires starting. The assassination and the striking of the match are events and both can serve as causes. Thus causal relations obtain between events, in which case one's ontology must, it appears, include events as well as bodies or substances.

If one believes that the striking of the match caused the fire, is it because there is some causal relation one is able to observe in that action? Other things may have immediately preceded the onset of fire. Perhaps a fly landed on the matchbook the moment before the fire began. One does not conclude that the landing of the fly was causally related to the fire. The reason that most people would cite is that they have seen flies land on matchbooks before, in similar conditions (of temperature, pressure, humidity, wind velocity, etc.) without fire starting. They have also seen matches struck on matchbook striking-surfaces many times, under similar conditions, and fire did follow the striking. So 'causality' is not identified in individual events but in repeated instances. And individual events do not recur again and again, only event-types do. It is because match-striking events are followed by fire-igniting events *with regularity* that the fire is attributed to the striking and not to the fly landing. Since causal relationships relate event-types rather than individual events, the latter are regarded as causally efficacious only derivatively and inferentially.

Some authors go a small step further and hold that causality does not inhere in the relationship of one event-type to another event-type but only in the elements of causal laws. That is, causality ultimately inheres in law-like relationships. For example, one may regard as a law-like generalisation a statement like the following 'Under conditions of low humidity, moderate temperature, etc., a match struck with pressure x p.s.i. against a match striking-surface in such and such conditions, will produce fire.'

Philosophical realists hold that there is some sort of real power that accounts for causal action, whereas most empiricists will deny any such thing or mode as 'necessity'. The view defended in Chapter 3 above (and Chernoff 2002) is that philosophical realism is plausible and attractive, but its extension to SR, and especially social SR, engenders serious difficulties and must be rejected.

What sorts of things are there in the social world that can stand in causal relations or can be related in events or event-types? Einstein, influenced by Spinoza, seemed to think that the physical world ultimately boils down to a single material substance. One might ask, in parallel fashion, whether the social world should ultimately boil down to a single social substance? But over the centuries physical theory has changed its ontology, from the point-particles in Newtonian physics to the continuous fields of the Maxwell–Lorentz theory and again to those of Einstein's theory. Is there a problem with these shifts? Does physical monism become a degenerating research programme, in Lakatos's sense? (See Fine 1996: 96.) The larger question, though, regards the change in physical theory that appears to bring it closer into line with the social world, namely, the shift to probabilistic laws and causation. How does this affect the notion of 'causation' in the physical and social sciences? The social world, just like the physical world, contains processes, events and event-types, which are the relata of causal statements. The conventionalist's commitment to event-types and events involving nation-states, institutions, etc., recognises that these help in the formulation of the best theories of IR. They account for the broadest range of phenomena, in the simplest way, with coherence and consistency. But one must remember that for all fallibilists a reorganisation of the data or the acquisition of new information could lead to the rejection of the currently accepted theory in favour of some alternative.

In any case, 'causality' is predicated of individual events, as both scholars and non-scholars ask for the causes of specific events. However, the argument that causality resides in event-types is quite compelling. While Hume may have unduly rejected the notion of 'causal powers', he is right in holding that one's knowledge of cause-and-effect relationships comes only from observing repeated instances of the event-types. There does not seem to be any other way that one could conclude that the lighting of the match is caused by its striking except by observing the constant conjunction – because other potential causal events candidates occur simultaneously with the striking on any individual lighting-instance. Only when many instances are observed may one conclude that the fly-landings, and other non-causes are not, while the striking is, constantly conjoined.[2] Those who argue that causal relations ultimately inhere in event-types agree that individual-level questions are bound to arise but can be answered only by reference to causal relations between event-types, and that constant conjunction, whether or not it is all there is to causality (and the causal realism endorsed here denies that), observation of constant conjunction is the basis, and a necessary condition, for knowledge of causal relationships. (The only possible exception would be effects of one's own decisions.)

Context-dependence

Claims about 'causation' in ordinary physical and social discourse are context-dependent. Typically, when one analyses the relationship of modern industrial powers and war, there is a tendency to ask for *the* (singular) cause of that relationship (e.g., many great power wars before 1945, none after). One must be careful not to give in to this tendency. Perhaps in some cases there is a single cause, understood as characterised below. But there are almost always many factors that play an important role. In some cases, while many factors are causally relevant, the question 'what caused x' is such that most of the causally relevant factors are assumed to be present and the mystery may be solved by identifying a single additional factor. When one looks at how the term is used in contemporary parlance and what would count as a 'right answer' to the question of 'What caused x?', there is a clear context-dependence.

Consider the outbreak of a fire. In one case an experiment is to be conducted in a vacuum chamber on a space shuttle. The apparatus has wires dangling inside the chamber, which occasionally become live. The past ninety-nine times experiments were performed in this chamber everything went smoothly. Since the vacuum chamber has no oxygen, the various loose, live wires were harmless. In the current case, on the 100th time the experiment is attempted in the vacuum chamber, it results in the unexpected outbreak of a fire. The fire-extinguishing system went into action and immediately put the fire out, but the experiment had to be aborted. The job at hand is to analyse the cause of the fire. Investigators pore over the well-preserved chamber, equipment and records. They discover a defect in a seal on the vacuum chamber, which allowed oxygen into the generally reliable vacuum chamber. The findings are: the cause of the fire was the presence of oxygen in the chamber.

What makes the presence of the oxygen the cause of the fire is that it is not expected. Under normal conditions, there is no oxygen in the chamber. In contrast, if one asks why a fire started in the engine compartment of an automobile, one might discover that gasoline leaked from a fuel line and reached the exhaust manifold, which was very hot and which ignited the gasoline. The cause of the fire was the presence of gasoline outside of the fuel lines. There was oxygen present in the engine compartment and its presence was just as necessary for the outbreak of the fire as in the space shuttle case. In the case of the engine fire, however, the presence of oxygen is not what one would normally call 'a cause' because one would expect it to be there; it would be quite surprising if oxygen was found to be wanting in the engine compartment. The high temperature of the exhaust manifold would not be regarded as a cause either, because it is normal for the manifold to be hot. In the context under consideration, the motorist driving her car, the exhaust manifold would become very hot. In both cases there was, and had to be, oxygen, spark, etc. But in the case of the space shuttle experiment, the presence of oxygen is considered the cause because in that context or setting, it is the element one would not normally expect to be present (as background). In the engine case it is the gasoline spilling outside of the fuel lines, because in that setting it is not what one would expect to be present (as background).[3]

A related distinction in IR is made by Waltz (1979), between active and passive causes. For Waltz passive causes are long-term conditions, such as anarchy, that one might expect to obtain under normal circumstances. The presence of anarchy is causally relevant for the behaviour observed in the international system. But it is rarely noteworthy in the investigation of individual cases, since it is always present. It does help explain the difference between cases in IR and cases in domestic politics, since it is normally present as a background condition only in the former.

Another related distinction between causes is that between 'proximate' and 'ultimate' causes. This differs from the distinction Waltz invokes in that one could imagine a proximate cause as either permissive or precipitating; furthermore, an 'ultimate' cause need not be a long-standing condition but could instead be an event (the ultimate cause, e.g., of dinosaurs' extinction or of the formation of the universe). Ultimate causes are universal statements that explain proximate causes. Proximate causes are deductively derivable from ultimate causes and focus on explanations of immediate occurrences.[4] Various other bifurcations of 'causation' are found in IR and social science discussions of 'causation'.

Probabilistic causation in physical and social science

The classical view of the physical sciences, as noted above, is that laws are deterministic and as such they differ from generalisations in the social sciences, which are, at best, probabilistic. For example, in most cases when a nation-state is threatened with attack by a coalition it forms a counter-balancing coalition. But in some cases the state capitulates and in some cases it defends itself unaided. Knowledge of the conditions does not yield a deterministic prediction of a defensive coalition. While the expected result may occur most of the time, it does not occur every time and cannot be guaranteed to occur.

Probabilistic associations contrast with the invariant, deterministic behaviour observable when a safe is dropped out of a window: its uniform acceleration is the same in every instance. Thus the obviously probabilistic nature of social relationships and the invariant regularities in the natural world led to the widely held view of a fundamental disanalogy between the natural and social sciences. While investigators searched for generations for a more physical-science-like foundation for the social sciences, an unexpected thing happened a century ago:

natural science laws were recast as bearing much more similarity to social science laws. The laws found now in physics, and apparently the nature of physical causality, became probabilistic.

Contemporary physics

Several developments in twentieth-century physical theory have pointed to the need for a probabilistic account of 'causation'. The correct understanding of quantum physics has, of course, been the subject of a well-known debate. Only a very small glimpse of it is possible here, which emphasises several central points of contention relevant to an understanding of social science theory. With regard to Einstein's view of modern physics it is worth noting that he maintained a Duhemian holistic perspective until the end of his life. Einstein argues at various points that the notion of 'truth', even the truth of a scientific proposition, only has meaning in the context of the theory in which it is embedded. Moreover, the theory can only be evaluated when conjoined with a mathematical framework. Fine (1996: 90) notes the constancy of this view from 1929 to 1949. Einstein (1949: 13) says:

> [a] proposition is correct if, within a logical system, it is deduced according to the accepted logical rules. A system has truth-value according to the certainty and completeness of its coordination-possibility to the totality of experience. A correct proposition borrows its 'truth' from the truth-value of a system to which it belongs.

The present study seeks to make sense of the notion of 'probabilistic causality'. It is helpful to understand Einstein's conception of 'physical causality'; he contrasts 'causal' (or as he prefers, 'non-probabilistic') laws with 'probabilistic' laws. (Einstein avoids the term 'probabilistic causality'.) Is nature probabilistic or deterministic? Even late in his career, Einstein still regarded the matter as one to be determined not by philosophical debate or *a priori* reasoning but by asking which theory works best with the available observations, that is, with 'coordination-possibility to the totality of experience'.[5]

Quantum theory, in Einstein's view, is unfinished and the statistical understanding of it is not fully articulated. There is more that can be said about the complete state of a system at a given moment than quantum theory offers (Fine 1996: 92). That is, its probabilistic claims do not exhaust all possible statements that can be offered about the actual state of a system. If the decay of an atom does occur at a specific time, then the probabilistic statement that quantum theory provides about the timing of the decay of that atom is incomplete.[6]

Many have interpreted Einstein, at least after the development of general relativity, as having been a philosophical realist, if not a scientific realist. The sort of philosophical realism Einstein considers combines the principles of causal-determinism and observer-independence. These seem to be the two chief features of philosophical realism for Einstein, which are causal determinism and observer-independence. Two other important but subordinate features are representation in a spatio-temporal manifold and a monistic ontology (Fine 1996: 105). The more standard definitions of 'philosophical realism', as well as the extension to SR, include a correspondence notion of truth, which Einstein does not endorse.

Einstein understands 'philosophical realism' in a fairly clear way, and in a way that differs from proponents of SR in contemporary philosophy of science. Einstein believes, for example, that the attempt to grasp reality was a goal of science but he does not say that any current theory should be taken as having successfully achieved that grasp. The search for reality is a *motivation* of the scientist, but a conclusion to that search should not be understood as characterising the state of science at any given moment. While Holton argues that Einstein was closer to a metaphysical realist, Fine offers a powerful argument based on Einstein's published and unpublished correspondence showing that he was far from holding this view and was, rather, a 'motivational realist' (Fine 1996: 109–11). In fact, Fine goes so far as to argue that Einstein was – in terms of the debate between scientific realists and anti-scientific realists discussed in Chapter 3 – actually closer to the anti-SR 'constructive empiricism' of Van Fraassen than to the scientific realist position of Boyd or the early Putnam (Fine 1996: 107–8).

Motivational realism can be clarified with a parallel to the moral desire to make the world a better place. It is beneficial if most scientists believe in the reality of theoretical entities postulated by the theories they endorse. But the value of the theories themselves stands apart from both the scientific desire and the moral motivation to push back the frontiers of knowledge or to create a scientific theory that improves life by better supporting technology of medicine, agriculture or pollution controls. This is similar to the way in which a desire to uncover the true nature of theoretical entities is a motivation that may stimulate scientific research in *some* cases, but is not a component of a scientific theory or of a foundational philosophical doctrine that justifies (even in a 'non-foundational' way) the scientific laws and theories.

Social explanation

The purpose of noting Einstein's view is that the account developed here is consistent with a 'motivational realism' of the sort attributed to Einstein. Constructive empiricism applied to the social science allows that there is a reality that underlies the observed phenomena. The actions and processes that the social sciences study are observable, within the limits laid out in the previous chapter. There is a real material world that constrains those actions, processes and structures. And there are real people, with real desires and beliefs that operate within the constraints of those material factors. As Wendt emphasises, the consequences of getting things wrong or getting them right are real enough for policy-makers and citizens, and getting them wrong can bring about serious real-world inconveniences. Lives are changed by policy decisions and the changes will be influenced by how good or how bad the predictions are; one's awareness of this enhances one's motivation to get the predictions right and choose appropriate policies that will move one closer to one's goals. The doctrine of motivational realism can stimulate one to seek greater understanding of the social world.

An account of 'causality' would then need to countenance 'probability'. How might this be done? One criterion of statistical causality that has been proposed is 'high probability' (Hempel 1965: esp. 381–405). If M is strongly associated with P, that is, if a sufficiently high proportion of things with the quality M are things with the quality P, then one might be tempted to say that M causes P. But there are straightforward problems with this approach for anyone who takes 'causality' seriously. For example, consider the McPeace hypothesis (MP) *if a pair of states both have McDonald's restaurants (M), they will (almost certainly) be at peace (P)*. The probability of the hypothesis MP is high, since data show that all cases of M are also cases of P. So it would seem that one may accept the proposition 'M causes P'. What about the jump from the high statistical association of M with P to the conclusion that M causes P? It does not look plausible.

Few would believe that the presence of McDonald's is enough to prevent states from going to war against one another. Rather, the plausibility of MP derives from the criteria that McDonald's use in awarding franchises. As a result, most McDonald's restaurants are in industrialised, wealthy countries, most of which have been democracies since McDonald's international presence began. The DP theory, or democratic-industrial peace theory, would account for the truth or high probability of MP. The hypothesis MP is not an acceptable causal hypothesis, despite the high association between M and P, because there is no reason to believe that Canada and the US, or Britain and France, purged of their McDonald's restaurants – but without any other changes – would be even slightly more likely to go to war with one another. As noted above, some authors get at this difference by arguing that MP fails as a causal claim because it does not support counterfactuals. IR authors who take this view include Cederman (1996), Jervis (1997), King et al. (1994), Patomäki (2002) and Russett (1996), among others.So high probability is not the key criterion of causality. In some cases low probability will suffice to indicate causal efficacy.

The key is that, however low the probability is, the probability estimate in the absence of the factor would be much lower still. A low probability, e.g., of 0.15, might also indicate causal efficacy. For example, suppose that for people exposed to industrial waste substance X there is a 0.15 probability of developing cancer type Y. The 0.15 value may seem low. But if the probability of that sort of disease among people not exposed to it is 0.01, then one might, despite the low 0.15 probability estimate, conclude that X causes cancer Y, since one's chances of that form of cancer increase fifteenfold when exposed to substance X. The regular association of X and cancer Y is the key to evaluating claims of probabilistic causation. There is no requirement of universal or constant conjunction. One further problem in the social sciences' use of probability statements is their sensitivity to how a problem is framed. Take for example Diehl's (1983) response to Wallace's (1982) analysis of the effects of arms competition. Diehl criticises Wallace's method of treating the Second World War as multiple cases rather than a single case, since Wallace codes each nation-dyad-year as a case. Thus Germany–Russia in 1941, Germany–US 1941 and Italy–US 1941 are three cases. Wallace considers the 1938–41 period in Europe as involving nine cases, and dyads involving Japan account for another six cases of arms escalations. Diehl notes that the correlates-of-war methodology codes a three-way dispute as a single case, while Wallace chooses to code each dyad as a case. Hence a three-way dispute, comprised of three dyads, becomes three cases. Thus where Wallace found a significant association between arms races and war, Diehl's method found no significant association. Less dramatic examples would not necessarily change an association into a non-association, but would move relative frequencies up or down the scale, depending on how the hypothesis is framed.

Probabilistic causality is important because the grand historical debate over whether a cause 'necessarily' follows its effect, which was the heart of Hume's criticism of the use of the usual notion of 'causality', can be ameliorated by recognising probabilities. The 'necessity' condition may be dropped in the probabilistic social sciences.

It should be clear that the probabilistic nature of causation does not raise difficulties for the naturalist, since Chapter 1 adopted the principle that expectations of 'knowledge' in the social sciences should not exceed any limitations found to constrain knowledge in the physical sciences. Furthermore, it should be clear that probabilistic statements are not all alike: different sorts should be interpreted in different ways. How then should the investigator treat the body of evidence from which the probability statement is generated?

Meanings of 'probabilistic causation'

The use of the notion of 'probabilistic causation' in both the natural and the social sciences raises the question of just what is meant by 'probability'. This section seeks to answer the question but points out that there is more than one answer, depending on the type of statement in question. Three points must be established, that there is a need for multiple interpretations of 'probability statements'; that the probability assignments of two statements may be the same, but they must still be treated differently because of different levels of confidence in the grounds on which they are accepted; and that there is a clear distinction between 'mature' and 'immature' sciences, which the notion of 'levels of confidence' helps to clarify.

Interpretations of the probability calculus

While probabilistic statements are central to the social sciences, much of social science methodology today, and IR meta-theory in particular, takes the interpretation of probability statements as unproblematic and as exogenous to the foundational account. This invites grave dangers, since there are genuine problems in understanding the appropriate meanings of probability statements, which have been interpreted in several quite distinct ways. A well-developed meta-theory should at least sketch a solution.[7]

The philosophical debate over probability theory has included at least four major lines of interpretation of probability statements. One is that they are logical relationships between statements, a partial version of the familiar notion of 'full entailment' (Carnap 1950). A second is that they are expressions of real propensities of entities and processes (Popper 1957). A third is that they are subjective statements about the beliefs of an individual investigator, that is, they show how strongly the believer in the probability statement believes the statement as an unqualified truth, which is typically discernible by the believer's behaviour relative to the belief (Bayes 1763, DiFinetti 1964, Jeffrey 1992). A fourth is that they are shorthand for statements about long-run relative frequencies (Von Mises 1939). In the case of the last interpretation, the reference class is often infinite or indefinitely large, and because dividing by an infinite quantity yields an infinite quotient, cases of the property in question will be equally frequent. As Von Mises points out, the relative frequency that is relevant is the *limit* of the frequency of a finite sample as the sample is increased indefinitely (Von Mises 1939, see also Reichenbach 1936). The relative frequency interpretation is empirical and factual, and it is often cited as the meaning scientists attach to statements in chemistry, quantum physics, genetics, etc.

Confidence levels

Some of the serious difficulties that arise in applying the account of probabilistic statements to the needs of policy-makers require further distinctions. How much can a policy-maker rely on a theory or law to justify a prediction regarding a policy's leading to satisfying a goal? A policy-maker has two decision problems. The first has policy options a, b and c and the second options d, e and f. The probability of policy a satisfying the goal of the first problem is 0.6, for b it is 0.3 and for c it is 0.1; these assignments are based on theory T1. The probability of policy d satisfying the goal of the second problem is 0.6, for e it is 0.3 and for f it is 0.1; these assignments are based on theory T2. The policy-maker chooses policy a for the first problem and b for the second problem. But neither is certain to lead to the relevant goal. How much effort should be put into preparing for the possibility that policies a and d will fail? If the two goals are of equal significance, does it follow that a consistent, systematic policy-maker should hedge bets equally in the cases of a and d? The answer is no, because he or she may have different confidence in the grounds for believing those estimates, namely T1 and T2. The policy-maker may need to hedge bets more in the case of d than in the case of a because he or she has less confidence in T1 than in T2. Thus the confidence the policy-maker has in a statement is different from the probability value assigned to that statement.

To clarify, consider an example in which one assigns a low probability, say 1/10,000, to a proposition like 'Italy and France will go to war next year' or 'It will snow in Los Angeles next summer', but attaches very high confidence to that probability assignment. On the other hand, one may assign a high probability value, 12/13, to a proposition like 'The next card drawn from the deck will not be an ace' but attach a low confidence value to it, as a result of the nature of the evidence available, e.g., that the investigator is in a room is full of professional magicians. In general, one would place higher confidence in observation statements than observation-generalisations, theories, or predictions.

Predictions will typically be justified by theories. Theories and theoretical laws are selected in part on the basis of accepted observation laws. Observation laws are largely derived from specific observations. A lowering of the confidence assignment would most often be a function of one's beliefs that the conditions under which the observation was made are 'normal'. Confidence levels will vary for all four types of statements, but there is a quasi-foundational ordering, beginning with individual observations. There is no strict ordering, because on a pragmatic theory of knowledge, any accepted statement can have an effect on the evaluation of any other statement. In a fallibilist theory of knowledge, all empirical propositions are subject to revision and rejection at a later time. The observation 'two ants are walking side by side' may be rejected once the investigator learns that there is a mirror along the bottom of the wall where the observation was made. On a fallibilist account all beliefs are suspect, but they are not all equally suspect. Observation statements like 'Montmorency is a dog' or 'Russia and Japan are not at war' are less likely to be rejected as scientific enquiry proceeds than observation laws like 'Liberal democracies do not fight one another.' And the latter are less likely to be rejected in the future than social science theories (like the functional theory of regimes or offensive realism) or than predictions derived from theories.

The goal of philosophical debate on probability theory has been a fully adequate unique interpretation of the 'probability'. However, the different types of statements that appear in social-scientific theorising cannot all be satisfactorily understood with a single interpretation. Carnap (1950) offers a very important insight, which is that there are distinct and irreconcilable meanings that speakers attach to probability statements, especially, in his view, 'relative frequency' and 'partial entailment'. He argues that these simply constitute two different concepts, both of which are intended by the speaker on different occasions and both of which serve useful functions. The present study takes a similar approach in that it endorses the idea that 'probability' has two distinct meanings, in different contexts: the relative frequency, and the subjective degree of belief. Observation statements that are used as evidence as well as predictions are accepted based on a subjective notion of 'probability', while observational laws and theoretical laws and theories are accepted based on a relative frequency notion.

Finally, it is worth pointing out that different confidence levels are not only different when one compares IR theories like DP theory to functional theory of regimes (Keohane 1984) but also when one compares DP theory to kinetic theory of gases or quantum theory. There are different confidence levels, different subjective levels of belief in the likelihood that DP theory (e.g., Russett and Oneal 2001) will some day be discredited than that relativity theory will be discredited. This should be recognised as a significant effect of different levels of maturity of different sciences, noted by Kuhn (1962).

Conventionalism

Conventionalism, probability and 'truth'

A critic of the conventionalist philosophy of social sciences developed here might be suspicious of its ability to help answer the central question of the value of IR theory for policy-making purposes, charging that conventionalism would be likely to dismiss prediction in the social sciences. Such a fear would note that conventionalism might overstate the differences between the social and natural sciences, since the latter are taken to be objective and predictive, while the former deal with the behaviour of humans, which might be held to be guided by convention. These conventions, furthermore, are often arbitrary, governing, for example the order in which states are seated in the UN General Assembly or on which side of the road motorists drive. But this dichotomy between the natural and social sciences requires a causal-deterministic view of the natural sciences, which is hard to maintain in the post-quantum theory world. The universality associated with 'necessity' can no longer distinguish the 'causal-deterministic' natural sciences from the 'merely probabilistic' social sciences.

The concepts and constructs of the social sciences are created because they serve an analytical and theoretical purpose. They allow the formation of hypotheses and theories that order experience and they provide the basis for explanation and action, since action requires prediction or forecasting. This fits with the claim, discussed in Chapter 3, that Cartwright makes to the effect that laws are idealisations and approximations. But it does not fit with the scientific realist's conception of the structure and ontological commitments of scientific theories.

Consider the social behaviour of studying IR theories. As one studies them, one categorises them. Are the categories given by the 'real social nature' of the theories (in a way parallel to that in which all physicists agree that the physical elements fit into 109 known non-human-made 'natural kinds' or that in which human individuals fall into one 'natural kind')? If so, it would seem that all IR theorists would agree on how to categorise them, e.g., on how many categories there should be. But that is not the case. Categories are chosen for the convenience of the author with respect to the nature of the study undertaken. For example, why does Waltz (1959) select three levels of analysis, while Hollis and Smith (1991) select four? They are each trying to categorise many existing theories. Why not choose as many categories of levels as possible? (Rosenau 1980: ch. 6 identifies six.)

The reason authors do not all choose the same number of categories is that the gain in simplicity of fewer categories counterbalances the gain in 'accuracy', in the sense that the cells of the typology are supposed to include 'like' theories and the theories within each cell exhibit much less variation when a dozen 'levels' or cells are included. Three was optimal, at least in Waltz's view. Because his *Man, the State and War* was published in 1959, it preceded the explosion of bureaucratic politics theories of foreign policy in the 1960s and 1970s. There were bureaucratic politics theories well before Waltz's book appeared, but they were not numerous enough to justify the addition of another 'level' in light of the loss of simplicity. The increase in the number of proponents of such theories is probably why, for Hollis and Smith in 1991, the loss in simplicity of increasing the number to four was considered worthwhile. For Hollis and Smith, staying with three levels would do too much violence to the accuracy of representing the scholarly debate. Similarly, as IR theorists choose what categories they should focus on with respect to agents in the world of international politics, they do not all choose the same agents or same categories of agents. Conventionalist foundations, as will be seen, help to illuminate why these differences are rational and part of good scientific method.

Conventionalism versus instrumentalism

Conventionalism is sometimes associated with instrumentalism. But the two doctrines have quite different implications. Conventionalism acknowledges two points that do not pertain to instrumentalism. First, false propositions may have true consequences. The rules of logic require the *salve vertitate* principle: if one begins with true statements, then any logical consequences of them must also be true. This 'truth-preserving' feature of valid inference does not entail any 'falsity-preserving' counterpart. If P is true and Q is false, then a truth table for P&Q will show that the conjunction as a whole is false. However, P&Q entails P. Thus, formal logic leaves no doubt that a false proposition may entail a true proposition (and indeed any number of true propositions. 'Russia is landlocked' is a false proposition. But it entails the true proposition, 'Russia is landlocked or Russia has a navy.' Hence, a theory that conjoins many propositions, some of which are false, must be regarded as false, if it has a truth-value at all.

The best theory that is available may contain statements that are in reality false, though they are not yet known to be false. If such a theory is taken as having a truth value, it would be false. However, conventionalists realise that such a theory still may have value because it allows the investigator to predict many future events that will in fact occur. That is, the predictive consequences are often true. Instrumentalism, on the other hand, eschews talk of truth-values from theories and may fail to see that some false propositions have true consequences and that some true propositions are true but not provable given the available evidence. In this regard, conventionalism is a logically superior account of science. So even an opponent of conventionalism, like Lakatos, is able to say, 'Conventionalism, as here defined, is a philosophically sound position' (1978: 106).

Poincaré

This study holds that conventionalism is capable of solving some of the key foundational problems noted above. The view adopted here is not a thoroughgoing empiricism that denies genuine causality in the social sciences. Conventionalism as a foundation for science was developed at the beginning of the twentieth century, during the (last moments of) the reign of classical mechanics. Poincaré's 1902 work *Science and Hypothesis* (1952) is often regarded as the major exposition of the doctrine. Although initially appealing, the doctrine soon lost support as a result of a number of objections. While conventionalism is often identified with the work of Poincaré, it will be argued here that the conventionalism of Pierre Duhem is preferable in a number of respects.

Poincaré is generally understood to hold that scientific theories are analytic and that the theoretical principles are true by definition. He does this in part by distinguishing between 'theoretical laws' (which Poincaré referred to as 'principles') and 'experimental laws'. Even after the discoveries by Lobachevskii, Riemann and Bolyai of various non-Euclidean geometries, the appeal of Euclid remained powerful, which is helpful in illuminating Poincaré's application of conventionalism to the physical sciences. Recall that when the twentieth century began, Euclidean geometry had been accepted for 2200 years as the 'true' geometry of space and its truth seemed beyond doubt – especially in light of its apparent re-confirmation in the seventeenth century by the astonishing success of Euclidean-based Newtonian mechanics. In the wake of Newton, Kant presented a highly influential and widely accepted argument according to which the truths of Euclidean geometry like, 'the shortest distance between two points is a straight line', are synthetic *a priori*. They are *a priori* because no experience can cast doubt on their validity. And they are synthetic because the concept of the predicate goes beyond the concept of the subject. Kant says in the *Prolegomena* that the notion of 'straightness' contains only quality and not quantity, so 'a straight line is the shortest distance between two points' must be synthetic (Kant 1950: 16). Even after the advent of relativity theory, philosophers of science had difficulty shedding Euclidean geometry. They offered a variety of arguments designed to retain at least some elements of Kant's foundations for geometry. A sign of the difficulty of giving up Kant is evident in Russell's 1897 *Essay on the Foundations of Geometry*, in which he argues that the geometry of qualitative relations of points, lines and planes (projective geometry) is *a priori*.

The third quarter of the nineteenth century saw the development and acceptance of non-Euclidean mathematics. Riemann developed systems of both constant and variable curvature. These systems were seen as non-contradictory and hence logically coherent, but were not applied to the geometry of space until Einstein, whose general theory of relativity (1916) first helped explain the Newtonian anomaly of the perihelion of Mercury. It was later confirmed by its correct prediction in the 1919 eclipse experiment (Gillies 1993: 85).

Poincaré, who was intimately familiar with non-Euclidean geometries, rejected Kant's *a priori* foundations of geometry. But he also rejected empiricism, since the ideal evidence on which geometrical systems are based are not obtainable from experiments or measurements on material objects. For Poincaré geometry is a system of measurement. The imperial system of the inch, foot, yard and mile is a system of measurement distinct from the metric system. For Poincaré the different systems of geometry are comparable to the different systems of measuring distance. Thus the question, 'Which system of measurement is the true system?' has no answer. No one would say the metric system is the 'true system' any more than the alternatives. The metric system supplanted the imperial system in most of the world because it is more convenient, not because it is truer. Similarly, for Poincaré, and contrary to Kant, no system of geometry is 'true'. Euclidean geometry is the most convenient. But it is true only in virtue of the system of definitions and stipulations that render its theorems acceptable. Poincaré says, 'If geometry were an experimental science, it would not be an exact science. It would be subjected to continual revision. Nay, it would from that day forth be proved to be erroneous, for we know that no rigorously invariable solid exists. *The geometrical axioms are therefore neither synthetic* a priori *nor experimental facts*. They are conventions' (Poincaré 1952: 50).

Poincaré sees the status of theoretical laws of physics as much the same as that of the axioms of geometry; they are neither *a priori* truths nor experimentally demonstrated. Newton's law of inertia is not *a priori*, evident from the fact that Aristotle denied it and proposed an incompatible law (of circular motion) that remained accepted in science for nearly 2000 years. Likewise, the law of inertia is not an experimental law, since the claim that 'a body acted on by no forces will continue in straight line motion' cannot be derived from experiment, since there are no experiments conducted upon bodies that are acted on by no forces (Poincaré 1952: 90). An experimental law can always be refined and replaced by a more accurate experimental law, in light of new observations. But Poincaré says that no one believes that such is the case with the law of inertia.

Observation may conflict with theoretical prediction. But in such cases, one may always add further postulates or alter existing ones to reconcile an old theory with experimental results (Poincaré 1952: 96). In view of Newton's theoretical laws, Poincaré says, 'the principles of dynamics appeared to us first as experimental truths but we have been compelled to use them as definitions' (1952: 104). Experiments 'will never invalidate' the principles of mechanics (1952: 105). 'Principles are conventions and definitions in disguise. They are, however, deduced from experimental laws, and these laws have, so to speak, been erected into principles to which our mind attributes an absolute value' (1952: 138). But only an unacceptable nominalism would hold that this exhausts all of science. Poincaré says that this shows that the English attempt to make mechanics experimental is correct, in contrast to the continental view, where it is treated 'always more or less as a deductive and *a priori* science' (1952: 89). Conventionalism was extended from geometry and mechanics to all branches of science, not by Poincaré himself, but by Edouard Le Roy (1901). Poincaré held that science has some experimental and some conventional laws. Only the latter are immune to refutation by experience. Le Roy offers as an example the principle that freely falling bodies accelerate uniformly. He says that the meaning of 'free falling' includes the notion of 'uniform acceleration'. An investigator who finds a falling body that does not accelerate uniformly would conclude that there are other forces interfering with the motion of the body. Hence the principle is true by definition (Le Roy 1901: 143–4, cited by Duhem 1954: 209). It is what Popper would classify as unfalsifiable.

The developments in the years immediately succeeding the publication of *Science and Hypothesis* led, of course, to a re-thinking of Poincaré's views. The special theory of relativity of 1905 undermined the view that the law of inertia is *a priori*. And a decade later general relativity undermined the idea that the space is Euclidean or that Euclidean geometry, because of its simplicity, would always be preferred and incorporated into science's best theories. There is controversy in Poincaré scholarship over how discontinuous the views of 1908 were, as presented in *Science and Method*. In general the conventionalism associated with Poincaré was seen as discredited by the scientific rejection of Newtonian principles (like that of inertia). Duhem, a part of whose position is defended here, offers a substantially different sort of conventionalism that is not affected by the acceptance of special or general relativity. While Duhem is widely classified as a conventionalist, his position is sufficiently at odds with Poincaré's that some, like Gillies (1993) do not regard it as properly labelled 'conventionalist'.

Duhem

Duhem's conventionalism stresses the holistic and the fallibilistic nature of scientific enquiry. With respect to holism, he argues that a scientific theory is a unified whole and it is impossible to treat a single hypothesis in isolation from the totality of principles, background assumptions and beliefs, e.g., about the accuracy of measuring instruments, etc. He says that a comparison of evidence to hypothesis is a comparison 'between the *whole* of the theory and the *whole* of the experimental facts' (1954: 208). With respect to Duhem's fallibilism, he states that, however good a theory appears on the basis of current research, it is possible that some as-yet-undiscovered theory will prove to be superior. In contrast to Poincaré, who held that some scientific principles cannot be shown to be false, for Duhem, falsifying evidence is *always* a possibility. But what is falsified is the holistic conjunction of hypothesis, theory, background assumptions and methodology.

In at least two senses Duhem denies the notion of the 'crucial experiment'. First, one might argue that a crucial experiment could provide the true theory or hypothesis if the investigator enumerates all of the possible explanations for the known phenomena and then devises and conducts an experiment that will reveal the truth of one and the falsity of the rest. Duhem would point out that even if one of the hypotheses correctly predicts the experimental result, it is impossible to attach certainty to the hypothesis because one can never be sure that all conceivable hypotheses (or theories) were included in the list enumerated. It is always possible that there will be a new hypothesis proposed that explains all current (or future) observations more fully. A second sense of 'crucial experiment' suggests that, given two theories T1 and T2, an experiment can be devised that will allow the investigator to disprove one and as a result show that the alternative is, if not indubitably true, at least superior. But Duhem, in one of his most important contributions to the philosophy of science, argues that scientific practice is holistic in that individual hypotheses or theories cannot be tested in isolation from a range of background assumptions.[8]

It is impossible to conclude that observation e refutes hypothesis h1, because h1 cannot be tested in isolation from other hypotheses and beliefs. In order to test Newton's law of gravity one must make assumptions about the laws of motion. Newton's law of gravity cannot be tested in isolation from other hypotheses and beliefs. So if an experiment to test it is conducted and the result appears to disprove it, it is not legitimate to conclude that the law of gravity is false. This is because, according to Duhem, it is also possible that there is a falsehood among one of the other background assumptions or laws of motion, all of which were required from the chain of inferences leading to the contradiction between the predicted outcome and the observed outcome. So all that one may legitimately conclude is that the conjunction of the law of gravity along with the laws of motion and all of the background beliefs (e.g., about the accuracy of measuring instruments, etc.) is false.

Duhem denied crucial status even in the case of one of the most famous 'crucial experiments' in the history of science, namely, Foucault's experiment designed to show whether the wave or particle hypothesis of light was correct. Light, if it were composed of waves, would travel through water more slowly than through a vacuum and, if composed of particles, would travel more rapidly. The experiment was ultimately seen by proponents of both hypotheses as showing the superiority of the wave hypothesis (although for a time supporters of the corpuscular hypothesis sought to repair it). Duhem argued that those who interpreted Foucault's experiment as definitive were incorrect. In order to carry out the process of conducting the test and reasoning through from the observations to the conclusion about which of the two hypotheses was refuted, the investigator needs a substantial store of auxiliary assumptions. Duhem concluded that the experiment could not decide between two hypotheses about the nature of light but rather 'between two sets of theories, either of which has to be taken as a whole, i.e., between two entire systems, Newton's optics and Huygens's optics' (Duhem 1954: 189).

Einstein published the special theory of relativity just after the publication of the major works of Poincaré and Duhem. The rejection of Newtonian mechanics that followed from Einstein had very different effects on Poincaré's and Duhem's theories. Poincaré, as noted, had claimed that some of Newton's principles were disguised definitions that need not (and would not) ever be given up (because they are the simplest conventions) no matter what the experimental results were. Duhem defended the apparently-similar claim that logic cannot force the investigator to relinquish a favoured hypothesis. The big difference between the two physicists is that Duhem did not leave the matter there. Good scientific practice will force the rejection of some hypotheses, since the hypothesis that appeared to predict incorrectly is part of a whole (i.e., system of hypotheses and principles) and enough counter-evidence could bring down the system 'under the weight of the contradictions inflicted by reality on the consequences of this system taken as a whole' (1954: 216). (Duhem warned that scientists should be alert to the fact that principles that they may accept without question may some day be overthrown.) Duhem's meta-theory clearly allows for the rejection of Newton and the replacement of his theory by a radically different non-Euclidean system, while Poincaré's does not. However, it was Poincaré who quickly embraced Einstein, while Duhem did not. Duhem, in a work composed while France was fighting Germany in the First World War, disparaged Einstein's work (in a way that would be ironic after the rise of the Nazis) as an aberration of 'the German mind' with its characteristic 'disrespect for reality' (Gillies 1993: 105).

As noted, when the hypothesis h1 produces prediction p, and an experiment shows not-p, what has been refuted is h1 *conjoined with* all of the theoretical principles and auxiliary hypotheses, h2, … hn, which were needed to conduct the experiment and to reason from not-p to not-h1. So what is really disproved is the conjunction h1 & h2 & … hn. It may well be that some of them are true. Which member of the set should then be rejected? Does a conventionalist answer that it makes no difference which is rejected because the choice is conventional – and ultimately arbitrary? It is important to note that Duhem does not answer in this way. He does not hold that conventionalism entails arbitrary decisions on these important matters. Rather, for Duhem the answer comes from '*le bon sens*' of the scientific community. In the case of the controversy over the nature of light, the corpuscular school could have continued to alter auxiliary hypotheses and patch up the underlying theory indefinitely. But the 'good sense' of mankind, or at least of the scientific community, would prevent that. Thus while science is, in an important sense, conventional for Duhem, the grounds for theory choice are not arbitrary. This raises the need for clear, rationally based criteria of theory choice (see pp. 25–7 and pp. 78–80).

Explanation, convention and social causation

Explanation

It is essential to bear in mind the relationship of causal laws to theories. Many reflectivist critics of naturalism attack the parallel of social and natural sciences by focusing on causal laws and arguing that they have no legitimacy in IR. At several points in his recent book Patomäki dismisses causal generalisations in a sentence or two, pointing out that in open systems associations can give rise to faulty conclusions about causality. Patomäki argues that the understanding of 'causality' that involves invariances or universal regularities 'can be easily shown' to presuppose closed systems, which are almost entirely absent in the real world of natural and social phenomena (Patomäki 2002: 76, 135). For example, one might examine the morning paper for the president's daily schedule and then, each day, note that the president lunched at the place noted in the paper. Because of the constant conjunction one might be inclined to conclude that the listing of the schedule in the newspaper caused the president to lunch at the specified time and place. This is an erroneous conclusion. However, there is more to the matter than the fact that sometimes such conclusions are false.

The temporal association between the listing and the lunch is, of course, not definitive, since further investigation will reveal that both the listing and the luncheons are effects of some prior, common causal factor, such as the decisions of the president's staff. One of the grounds for endorsing causal laws as part of a proper theory is that the problem of understanding the connections between events or types of events becomes less severe with the inclusion of causal explanation; an explanation of how the one causes the other will reduce the danger that this sort of mistaken causal relationship might be embraced.

Cognisance of the association alone can still be of practical value, even if one does not understand the causal mechanisms or have access to a compelling explanation. Without knowing much about the causal mechanism it is beneficial to be aware of an association between the listing on a schedule and the president's luncheon, between the speedometer reading 200 k.p.h. and the police issuing a traffic citation. One must recognise that the association lacking a causal explanation is much more likely to be overthrown upon further investigation, hence less risk should be run based on the belief that the connection will continue to hold indefinitely.

Jervis (1976: ch. 6, esp. 227–39) argues that policy-makers have a tendency to draw lessons from historical events – but they are the wrong lessons because they do not seek the underlying causal principles, which is to say, they do not grasp the underlying causal relationships. Policy-makers carelessly detach the conclusion from the body of evidence that justified it and then reapply the conclusion in inappropriate circumstances. The solution is not to purge all causal laws as potentially leading to false conclusions, but rather to remain aware of the fallible nature of empirical disciplines including IR and to continue to revise and improve theories that policy-makers use.[9]

The view that IR meta-theory should be tied to empirical testing and the defence here of causal realism (discussed in the next sub-section) entail that reflectivists are wrong to deny causal generalisations. There are real causal connections in the social world. Furthermore, the criterion of empirical testing would strongly suggest that reflectivists are mistaken in dismissing causal generalisations. Admittedly, IR has a long way to go in developing theories which contain general (probabilistic rather than universal) propositions that receive discipline-wide consensus. Still, it has advanced far enough to have achieved some success, which constructivist and reflectivist views reject. For example, consider the statements 'Democratic and industrialised states fight one another less frequently than states generally because of their democratic and industrialised nature' and 'The military strength of militarily stronger coalitions causes the defeat of militarily weaker coalitions in war.' The rejection of the possibility of causal generalisations because of the 'open' nature of the international system and the possibility of erroneous conclusions would seem a weak reason to prohibit the IR theorist from accepting these generalisations.

Causal mechanisms and causal realism

Daniel Little (1993a) takes up Cartwright's distinction between two types of laws and applies it to the problem of prediction in the social sciences. He uses slightly different terminology, contrasting 'governing and phenomenal' regularities rather than 'fundamental and phenomenological', as Cartwright does.[10] He argues that in the social sciences there are phenomenal regularities but there are no governing regularities.[11]

Little contends that 'causal explanation is at the core of much social research and causal hypotheses depend on appropriate standards of empirical confirmation for their acceptability' (1993a: 185). King et al. (1994: 75) seem to echo this sentiment.[12] Little defends 'causal realism', according to which, there are many types of social cause; causal explanation is the central form of social explanation; causal relations are constituted by causal powers of various social entities not by regularities or laws; and micro-foundations are part of complete causal explanation (Little 1998: 197–8). For Little, the social sciences should aim to discover mechanisms that are derived from agents and institutions, which in turn produce regularities. Causal mechanism is 'a series of events governed by lawlike regularities that lead from the explanans to the explanandum' (1991: 15). Thus one should seek to discover the causal mechanisms and underlying causal properties, rather than the regularities that they produce (1993a: 185). In Wendt's references to Little (1991) he holds that Little is a scientific realist with regard to the natural sciences and an anti-realist with regard to the social sciences because Little does not accept the mind-independence of the objects of social enquiry (Wendt 1999: 71).[13]

Little's causal realism resembles the position of Cartwright, who also takes causal explanations as real. But it differs from Cartwright's in a significant way, especially if Little thinks that the primacy of causal explanation holds also for physical theory: Cartwright regards mathematical laws to be just as explanatory as causal assertions but does not regard the former as causal. She says that when one thinks about explaining phenomenal laws via mathematical equations, one must acknowledge that the former do not 'bring about' the latter. Thus Cartwright countenances an important class of natural scientific explanation, which may reflect fundamental/governing regularities, but which she does not regard as causal. In contrast, according to Little's account of the social sciences, all explanation is causal.[14]

Because social regularities are weak, one must pay more attention to the specifics of the social and individual-level mechanisms that produce the regularities and to the exceptions to the regularities. The search for laws is thus not the appropriate task of social science. Rather, given Little's endorsement of causal realism (1991, 1993a, 1998) and his view that causal analysis and identification of the underlying causal processes, not subsumption under general laws, should be seen as the core of social explanation, investigators should focus on the discovery of mechanisms that derive from agents and institutions, which in turn produce the regularities. One can adduce empirical support for the theories of the underlying causal regularities. So Little prefers social science to be directed to these micro-level investigations, e.g., those typical of the micro-foundational accounts drawn from rational choice, theory of institutions, collective action, game theory, micro-economics and micro-sociology.

Little argues that when it comes to causation, the power of institutions appears only through the powers of individual agents. 'Institutions have effects on individual behavior (incentives, constraints, indoctrination, preference formation), which in turn produce aggregate social outcomes' (1993a: 195). The power of the US Federal Reserve Bank to alter the environment in which choices are made 'is derivative on facts about typical consumers'. Hence, '[s]ocial entities possess causal powers only in a weak and derivative sense' (1993a: 195). These causal powers do not depend on the existence of law-governed regularities. The rock-bottom causal story is about the characteristics of typical human agents.[15]

Little urges that investigators should combat the impulse to use regularities in social explanation, which he thinks comes from the influence of the covering law model of explanation and from an uncritical naturalism, i.e., from 'an unhelpful analogy with the natural sciences' (1993a: 184). One should look for the underlying causal properties and discover them, rather than the regularities that they produce (see Jervis 1976). Little does not ultimately deny that a 'science' of the social world is possible. The regularities of social science are law-like in the crucial sense that they support counter-factuals. But the regularities are phenomenal: '[t]he causal properties of social institutions and the micromechanisms that underlie them, give rise to phenomenal laws, and these are the chief regularities identified by social scientists – not governing regularities' (Little 1993a: 197, see also 1993b: 364, 1998: 249). Little concludes that social science is capable of being genuinely 'scientific', but lacks natural-science-like governing regularities.

Causal mechanisms are captured in governing regularities, which are associated with theoretical laws (discussed in pp. 65–6, 77–8) and the latter involve terms that are required to provide causal explanations for the associations cited in phenomenal or observation laws. As noted, observation laws provide some confidence in the association. But one must treat them tentatively until they are 'understood' better within a theoretical framework. There is a greater danger of spurious or epiphenomenal relationships, as was seen in the McPeace hypothesis, MP. Even a 'bad' causal explanation is better than no causal explanation, especially if one recognises its tentative and fallible nature, since such explanations can contribute to progress. That is, they can serve as the basis for refinements of causal explanations as they are integrated into the theoretical and evidential framework. Part of what makes bad explanations bad is that they insufficiently cohere with the rest of the theoretical corpus (e.g., the hypothesis MP). For example, in social science explanations part of what would render an explanation bad is that it is not fully consistent with the rest of what the investigator has justifiably inferred about the social world or about human behaviour. As the explanation is scrutinised, any contradictions will emerge and it may be possible to modify them to produce better causal explanations.[16] The causal realism of Little is very appealing. Little is right to hold that there are causal processes and mechanisms, as he defines them, and in claiming that they are ultimately grounded ultimately in human agency. However, the evidence for the existence of a causal connection is a set of regularities that are detectable through observation. Little seems to underestimate the latter.

Causal mechanisms in observational but not theoretical laws

This study endorses common-sense realism but not SR, in part because the latter incurs the cost of introducing false propositions into the accepted body of beliefs without any off-setting benefit. The distinction is central to understanding the role of causality in causal conventionalism. Consider a researcher proposing an observed association but admitting that there is no plausible causal mechanism to explain it. Some empiricist philosophers would count such an association as theoretically appealing as any for which there are plausible causal mechanisms. The causal-conventionalist account, as the name indicates, is not purely empiricist in this regard and would not accept such a relationship as part of a properly constituted natural- or social-scientific theory. But at the same time, while the causal-conventionalist account does require that a theory should include specific causal mechanisms to explain observational associations, it does not require a theory to include specific causal mechanisms that connect theoretical entities. Although the theory does not endorse particular causal mechanisms, there must at least be 'plausible' causal mechanisms that may be hypothesised.

Conventionalism, in the forms that both Poincaré and Duhem advocate, eschews causality as part of scientific theorising. The causal-conventionalist (CC) account developed here endorses a role for causal relations and mechanisms and for explanations incorporating them. If CC rejects SR, then in what way can it endorse causality? The answer derives from CC's endorsement of common-sense realism. Scientific realists include causal claims but insist that theoretical entities be taken as real (that is the core of SR). The view of causality in the CC account developed here is different from these and flows in part from CC's endorsement of common-sense realism. Causal relationships hold in the physical world ('striking ceramic urns with cricket bats causes them to break') and in the social world ('spanking children causes crying', or 'in specified circumstances, surprise attack causes the attacked to retreat'). If one believes in causal connections in the social world, then the theoretical postulation of empirical social associations entails that there are causal connections between the events or event-types (surprise attacks and retreats) identified. If there are observed empirical associations that are incapable of explanation by means of any plausible causal mechanism, then the model, law or theory that makes use of the associations will not rise to the standards of 'acceptability'.

If one believes in the causal efficacy of social events or event-types, then one must reject any association that is incompatible with a plausible causal story. However, if a plausible causal connection is hypothesised between the theoretical entia, there are grounds for accepting the association. Nevertheless, the causal explanation for it may be erroneous. It should not be accepted as true in the way that the observable association itself is. Thus, the existence of a plausible causal mechanism shows the possibility of such mechanisms (*ab esse ad posse*). However, the postulated mechanism may not be correct. It may not capture the 'real causes'. Thus, plausible causal connections or stories are necessary for the association or law to be acceptable, since plausibility is not a sufficient ground for acceptance. There may be other causal stories that involve theoretical entities and these others may be true. Adding the danger of error into the theory by asserting that the theoretical causal connection is true is not counterbalanced by any practical or epistemic gain and is not a part of CC. The association will provide grounds for the pragmatist-oriented philosopher of social science to be able to formulate policies.

Common sense realism may be applied both to the physical and the social world. Those who endorse common-sense realism hold that 'striking the urn caused it to break' is a true statement under some circumstances (that is, if and only if striking the urn did in fact cause it to break). A supporter of common-sense realism may without any contradiction deny SR and so reject the 'real existence' of theoretical entities or at least the 'real existence' of any particular set.[17] Similarly, CC supports the view that specific causal claims consistent with common-sense realism may be endorsed but not purely theoretical ones. Common-sense realism and the distinction between 'observation' and 'theory' permit an analysis here according to which the 'non-epistemic seeing of phenomena' allows those phenomena to be taken as real. Theoretical entities are not. If one accepts Dretske's argument outlined in Chapter 3, then one may apply his distinction between 'epistemic seeing' and 'non-epistemic seeing' to social phenomena. Dretske himself uses 'seeing a battle' as an example. Thus, CC supports specific causal claims between event-types or events when the elements of the causal mechanism are observable. When theoretical entia are invoked, what is essential is that there be some plausible mechanism. However, no particular mechanism is accepted.

Convention and causation

One might object that the attempt to combine conventionalism with causal realism is just as inconsistent as Wendt's attempt to combine positivism with constructivism. But the two cases are not parallel. The conventional element in scientific theorising must be admitted regardless of whether or not one believes that theoretical terms refer, regardless of whether or not one accepts IBE as a warrantable form of reasoning, and regardless of whether or not it is a miracle that the scientific theory has grown more and more powerful.

Are *prima facie* reasonable desiderata always defensible? Consider someone not familiar with the experimental and theoretical literature in physics. Such a person might request of a theory that it tell him or her absolutely where in the universe – not merely their relative position – he or she is located at a given moment, even if the universe contains no matter beyond his or her body. A system of formal logic that is demonstrably consistent is preferable to one that is not. A system that is demonstrably complete (in the sense of accounting for all of the truths of elementary number theory) is preferable to one that is not. Imagine someone unfamiliar with the philosophy of mathematics, who longs for a system of formal logic that is both demonstrably internally consistent and demonstrably complete. These both seem to be reasonable things to seek, at least on the surface.

Neither the physical nor the mathematical desiderata can, upon analysis, be sustained. Both are known to be theoretically impossible. When one probes both available physical theory and data, one sees that there are no grounds to support any theory of absolute space or time, as Leibniz correctly argued three centuries ago, and as Einstein finally persuaded the scientific world.[18] And since 1931 it has been clear that the seemingly innocent desire for a demonstrably complete and consistent axiom system for elementary number theory is impossible to satisfy. Therefore, it is unreasonable to demand, or even hope for, a system able to demonstrate both features (Gödel 1931, Kleene 1952, Van Heijenoort 1966).

Similarly, someone persuaded of the value and plausibility of (at least some) causal propositions in the sciences might demand that scientific knowledge avoid all qualifications that make use of convention. However, an examination of physics shows clearly that there is no defensible way to demand any measurement of space that does not require conventions about the nature of the measuring instruments or the mathematics of measurement. Hence, if one is to endorse any physical causal claims, those claims will have to be combined with at least a limited conventionality, as formulated in the CS thesis that there is a conventional element to all scientific theory choice (introduced in Chapter 1, above and discussed further in Chapter 7 below). The persuasive core elements of Little's causal realist arguments (though modified in the argument of this chapter and the next) have considerable force.[19] But regardless of whether one argues for or against causality in sciences, one must accept the conventionality.

This study appends an analysis of 'causation' to a form of Duhemian conventionalism. While Duhem does not endorse 'causation' himself, there is no inconsistency in combining the two. And the procedure, in general, is one that others endorse. For example, Patomäki has added a causal component to theories that were developed without it. He adds a theory of 'causality' to the British institutionalist theory that he finds appealing on other grounds. He acknowledges that the theory was designed by its authors without a notion of 'cause'. As long as the theory does not have implicit features that prohibit the appending of a causal account, it is entirely permissible to so append an account, even if its originators might object. The same procedure is adopted here with respect to the quasi-Duhemian meta-theory defended. Duhem does not endorse a causal account and is in fact quite hostile to it. He offers reasons for denying causality in the physical world. But those reasons are logically distinct from those he offers in support of other features of his conventionalism.

Suganami's analysis of 'causation'

One of the most extensive and probing analyses of 'causation' in the IR literature is that of Hidemi Suganami (1996). In his excellent book on the causes of wars he examines and critiques how historians and IR theorists use the notion of 'cause'. Suganami also critiques some of the most well-known philosophical analyses. Suganami helps to dispel many long-standing confusions by his conceptual analysis of 'causation', by means of his separation of logical (or definitional) enquiry from scientific (i.e., empirical) enquiry, and by his careful distinction of different questions that scholars, at different times, ask about war (namely, what factors present in all wars? What factors are frequently associated with wars? And, why did some particular war begin?). The focus on logical analysis aids him in showing that there are logical prerequisites for war at all of Waltz's three levels of analysis and that various proposed solutions for war are logically incoherent. At the same time, he maintains that showing that something is logically necessary for war (like anarchy) cannot translate into a programme that has practical value.

Waltz and permissive causes

Suganami takes as his points of departure the most influential treatments, Waltz's analysis of 'international stability' and the causes of war, and Hempel's analysis of 'causation' (which Suganami calls, following Donagan 1964, the Hempel–Popper view). Suganami (1996) criticises Waltz's (1959) typology of 'causes' as 'efficient' and 'permissive' by arguing that Waltz includes everything that is not permissive into the 'efficient' category, which combines very disparate sorts of factors having little in common. In particular, Suganami shows that Waltz's argument (1959, 1979) that anarchy is a 'fundamental' cause (because it is permissive), which thus accounts for the recurrent nature of war, is inconsistent with his rejection of 'human nature' as a cause of war. Waltz argues that, by definition, 'human nature' is everywhere and always the same. Thus it cannot account both for the fact there was peace in Europe in 1910 and war in Europe in 1914. Suganami correctly points out that exactly the same can be said of 'international anarchy'. For Waltz 'human nature' is a causal factor, if at all, because it is an efficient or precipitating cause. But why must this be? Waltz argues that only third-image, i.e., international systemic, factors can qualify as permissive (and thereby fundamental) causes. First- and second-image factors, if they are causes at all, are efficient causes. Suganami's criticism here is well founded and goes to the heart of Waltz's argument that the third image is the most important level of analysis.

Suganami criticises Waltz for finding only a single 'underlying cause' of war, international anarchy. Suganami (1996: ch. 5) contends that some wars have other underlying causes. He argues that Waltz's account of the origins of particular wars is highly underdeveloped. But Suganami perhaps goes too far in the other direction (as will be seen in the next subsection) by focusing on the distinctive features of particular wars and by arguing that general claims or commonalities across cases are unimportant or inessential for causal analysis. Common features of different cases are essential in understanding causal forces.

Singular statements, regularities and evidentiary interaction

Suganami examines and rejects both Hempel's covering-law account of explanation and the regularity-oriented analysis of 'causation'. Certainly there are errors in the covering-law account. While there are many well-known criticisms of the covering-law model, Suganami's criticisms are not universally successful. The aspect of the covering-law model that he attacks is, however, an important part of the understanding of 'causation'. Suganami carefully distinguishes the problem of the meaning of a causal statement from that of the evidence needed to support such a statement. On the question of meaning, Suganami develops his argument by examining the claim that a covering law is part of any proper assertion of causality. He considers a stronger thesis, that one must know the covering law in order to advance any singular causal claim and a weaker thesis, that one may not always know the exact relevant covering law, but one knows that there is a relevant covering law whenever one asserts that there is a relation of causality.

According to Suganami:

> [w]e say, for example, that my having been in contact with someone who had a cold caused me to develop one even though we do not thereby claim to know what laws underlie the assertion … Similarly, historians offer causal accounts of the outbreaks of wars, even though they would not thereby claim to know what laws relating to the outbreak of war underlie their assertions.
>
> (1996: 120)

This example, however, is entirely unpersuasive, since people do believe that there is a cold-exposure and cold-transmission regularity; only because they believe in that regularity do they say that Lydia got a cold because of exposure to Patrick, who had a cold. There might be (and perhaps always will be) more unknown details about the mechanisms of cold transmission. But one thing that people do know, or believe, is that exposure creates a possibility for illness. In the absence of such a belief, no one would say that Lydia's exposure to the afflicted Patrick was a cause of Lydia's developing a cold.

Suganami says that the stronger thesis, that one must know causal laws in order to assert a singular causal statement, is not substantiated and that only the weaker version may be sustained. But what the weaker thesis, he says, 'entails *at best* is not laws but their existence' (1996: 123). Suganami's position here is, uncharacteristically, rather loosely phrased and consequently misleading. The argument he considers, contrary to his charge, does entail laws that have content; but they have less content than laws that offer more specificity, e.g., about causal mechanisms. 'Exposure to colds is a cause of colds' is a law but it has little content and no causal mechanism. It may be merely an explanation sketch, in Hempel's sense (1965), rather than a full and complete explanation. As more is learned, more detail is added to the explanation's laws.

Suganami's problem could have been solved by reference to Cartwright's discussion of the distinction between fundamental and phenomenal regularities. We may not know the fundamental law (or the detailed theoretical explanation in biology for the transmission of colds between humans) but we do know that there is a phenomenal regularity that connects cold-exposure to cold-catching. Imagine how an investigator would react to the claim that S1 developed a cold because of her exposure to S2, who had a cold at the time, if the investigator *did not* believe in the phenomenal regularity. In that case, the investigator would promptly reject the causal claim. Suganami says that at best the investigator knows only that there is regularity, but does not know the content of the regularity. But this is questionable, since the investigator, as just noted, would know the content in so far as it connects cold-exposure-type events with cold-transmission type events. There are other regularities he or she may not know, regarding details of the causal mechanism. But this outline of the causal mechanism involves a regularity whose content is known. A more plausible line of criticism for Suganami here would be the anti-naturalist one, denying that there is a parallel between someone's claims about the cause of a cold and an historian's claims about the causes of wars. But Suganami does not take this approach. Suganami does make reference to 'higher-level' laws (1996: 137–8) but does not distinguish lower and higher in a way parallel to empiricists and Cartwright or Daniel Little.

Deduction, according to Suganami, can justify a causal conclusion about a particular, unique case without recourse to regularities. When the investigator eliminates all of the other plausible candidates, the one that remains, no matter how unique, may justifiably be inferred to be the cause. On the surface this seems like a clever and water-tight argument that conforms to the practice of historians' analyses of wars. But below the surface it leaks badly. The historian might identify hundreds of factors or events, like a snowstorm in Antarctica, that occurred just before the onset of the war in question. How does one eliminate such irrelevant events from the list of possible causes and reduce the many down to 'likely candidates'? What is it that makes a candidate a 'likely' candidate?

Certainly one answer is that past experience shows an association between the factor and the onset of some other war. Thus there must be some plausible (probabilistic) generalisation involving that potential causal factor. An arms race between the two states initially involved in the war is a plausible candidate because there are past observations of arms races occurring prior to the onset of wars. (The existence of such generalisations leads investigators to seek mechanisms.) If there were no need to consider (often implicit) reference to generalisations, then it would be extraordinarily difficult to imagine how the historian would select a particular factor as a likely causal candidate. The plausibility of factor c as a likely candidate comes from the historian's knowledge of other cases. The past association between factor c and war need not be deterministic or necessary, that is, it may not apply in all known cases. It may be probabilistic and there may be alternative paths to war. Presumably the historian does a better job of identifying appropriate likely candidates and actual causes because he or she has more knowledge of other cases.

Suganami (1996: 151) states his view regarding the relationship of singular to general causal statements by saying '[a] causal explanation of a particular event … does not involve statement of the law(s) covering it. The ideographic does *not* presuppose the nomothetic.' This seems a correct statement of what is not the case. Still, one may wish to know what the relationship is. A persuasive case may be made that the ideographic and nomothetic are co-generative; advances in one helps us to see the other more clearly. When one has a plausible causal explanation for war x one may then apply that wisdom to arguments about wars in general. When one makes advances in our study of wars in general, then one may apply what has been learned to understand more deeply that perhaps factor b and not a, as previously believed, was the most important factor in bringing about war x.

The co-generativeness does engender a circularity. But it is a virtuous and not vicious circularity because it is presented here in the context of a meta-theory that eschews foundationalism in favour of fallibilism (as most contemporary meta-theories do). If there were an identifiable stopping point of analysis, this sort of circularity could present an epistemological problem. But in this context the discovery of causes in particular wars and in war in general mutually reinforce one another – or discredit previous general or particular hypotheses about war(s) – and in so doing make way for new analyses. The co-generation relationship may present difficulties for supporters of foundational theories of knowledge. But there is no such problem for those who endorse pragmatic or other fallibilist theories, since there is always a re-evaluation of existing relationships of confirmation as new evidence emerges. The co-generative thesis only implies that both new individual associations (e.g., 'factor x is believed to have caused outcome y under conditions z') provide grounds for reassessing the support of general hypotheses (of the sort 'factor x causes outcome y under conditions z') *and* vice versa.

Reasons, causes and explanation

Suganami does not accept the 'explanation–understanding' dichotomy in the way many interpretivists do. He holds that reasons are distinct from causes but, again, not in the way interpretivists assert. For Suganami reasons may be evaluated on moral grounds while causes may not. He cites the argument of Scriven (1959: 449) according to which an explanation is regarded as 'essentially a linkage of what we do not understand to what we do understand'. Suganami adds, '"explanation" necessarily involves "understanding" inasmuch as there can be no linkage, and hence no such thing as an explanation if we understand nothing'. He then cites Hanson (1958: 54) who says, 'We have had an explanation of x only when we can set it into an interlocking pattern of concepts about other things y and z.'

Suganami considers the anti-naturalist claims that natural science and social science explanations are fundamentally different, in particular because only the former conform to the covering-law model. He rejects this view, but not for the reasons that naturalists traditionally do (namely that social explanation conforms, too), but because neither conforms to the covering-law model of explanation. While Suganami is right in pointing out the inadequacy of the covering-law model, there are some correct aspects, which Suganami disparages, particularly the implicit reference to multiple cases (deterministic or probabilistic regularities) in singular statements about causation.

The main thrust of Suganami's analysis is consistent with the account proposed here. For Suganami, causation, explanation and narration are inextricably intertwined (1996: 150). Suganami's analysis fits with the fallibilism defended here, since he says, '[o]ur view suggests that the sequence of the events narrated … in such a way that the occurrence of the war can be made [more] intelligible [than before] constitutes the cause of the war' (Suganami 1996: 150–1). There is no inherent end-point of enquiry or certainty arrived at on such a view.[20] He comes out as a strong advocate of 'causation' in the social sciences, since he sees strong parallels between natural and social causation. Suganami (1996: 139) says that both social science and natural science explanations require a narrative account that makes the event intelligible. He sees causal explanation as a series of linked elements and, as fallibilists do, as never terminating in a once-and-for-all explanation. One main difference is that Suganami denies the importance of regularities in explicating or asserting 'causation'. The analysis presented here disagrees with one of the key claims of Suganami's view, which is that singular causal claims may be defended without explicit or implicit reference to general causal claims.

Reasons as causes

One of the important questions raised about causal analysis in the social sciences is the role of 'reasons'. Some authors deny that reasons are causes and hence that reasons can play a role in a causal analysis of international behaviour and social action. Two principal grounds are cited for this reluctance. First, cause-and-effect relationships obtain between events or event-types, and reasons are neither events nor event-types. And second, causality requires a contingent connection between 'cause' and 'effect', which the relationship of reason-to-action lacks.

There are some points where CC overlaps other meta-theories. As noted in Chapter 1, while CC strongly opposes the 'primacy of ontology' arguments of SR and critical realism, it agrees that there is a legitimate place for causal theorising in the social sciences and that reasons may be construed as causes. The grounds offered here for considering reasons as causes are, however, quite different from those offered by CRs like Patomäki (2002: 87–9) or Bhaskar (1986).

The considerations that follow combat these objections to understanding reasons as causes. The view defended here is that some reasons may properly be construed as causes. Some causes, e.g., physical causal relationships, are not reasons. And some reasons, e.g., the reason why there are no prime numbers between 61 and 67, are not causes.

Reasons and events

With respect to the first objection, that causation obtains between events or event-types, there is a pertinent distinction between two ways in which the term 'reason' is used. According to the first, a reason means 'S being in a state in which S has a reason for doing A'. It is an intention or goal of S (where S might be an individual, ministry, agency, national government, etc.) The expansion of national power is a reason, which does not appear able to serve in a causal explanation of India's testing of an atomic weapon or of the US's proceeding with missile defence tests. However, Prime Minister Vajpayee's goal of the expansion of Indian national power or President Bush's goal of the expansion of American power are different. Each is an example of a reason-state (Audi 1993: 234–5). While they are not events that involve change, they are conditions in the world. Thus their status seems to be akin to that of structural causes or what Waltz calls 'permissive' causes, such as anarchy.

While reason-states are not passive, both reason-states and permissive causes are long-standing and do not constitute sufficient conditions for the effect. Nevertheless, they play a causal role and have efficacy in conjunction with an active, precipitating cause (like Vajpayee's or Bush's election to office), which is ordinarily understood to be an event. Viewed in this way, the first objection to reasons as causes seems inconclusive, unless the field of IR theory is willing to reject entirely the view that structural factors may have causal status. This study endorses the view that standing conditions may properly be viewed as 'causal', which is advanced further in the succeeding section on 'context-dependence'.

The second objection to construing reasons as causes is that the connection between them is logical and not contingent. Analyses of 'causation' regard the connection between 'cause' and 'effect' as something that is to be found by observation and study of the empirical world; the two must be capable of being defined independently of one another. The relationship is thus regulative and not constitutive. Critics of counting reasons as part of causal analysis, especially those in the HT, argue that the analyst must determine what reason is at work before he or she is able to define the effect action; that is, the reason for action becomes part of the definition of the action itself. The alleged cause and effect cannot then be characterised or defined independently of one another. For example, one might identify one group of people dressed in blue uniforms and another in red uniforms discharging firearms at one another. Should this event be described as an attempt by the blue-clad person to rob the red-clad person, as a random homicide, an incident of gang violence, or war? An answer is needed in order to be able to state what the action is that the analyst is trying to explain. Many reflectivist and HT authors would argue against the conventionalist position defended here because they deny that reasons may be construed as causes. They hold that one can only define the action as war after the reason for the firing of the weapon is understood. Thus social causation is an unacceptable stretch of the ordinary use of the term 'cause'.

The independence of cause and effect and the argument from the origins of 'causality'

Is 'logical independence' truly essential to our basic notion of 'cause'? To answer this question it is necessary to have a fuller understanding of the meaning and origin of the term 'cause'. The objection states that the extension of 'cause' from the physical to the social world violates a basic feature of the ordinary understanding of 'cause'. However, the history of science in antiquity can be used to undercut that objection and to support, rather, the claim that the social notion of 'cause and effect' is more basic and pre-dates the natural scientific notion.

The standard view of the connection between the scientific notion of 'law' and social behaviour can be made by considering the traditional explanation of the origin of the notion of 'causality'. According to that account, people have always been aware of the connection between their willingness to perform an action and their performing that action in such a way that (i) the will precedes the action, (ii) the will brings about the action and (iii) the will does so necessarily. On this view such an agent's experience of connection between will and action was the precursor to the scientific ideas of 'causation' and 'causal law'. Hence, when a natural event occurs, it must do so with purposes similar to those that drive human actions. There must be a personalistic cause (Zeus disgruntled) that accounts for the observable natural phenomena (lightning bolts). This conception is helpful for those who defend the attempted extension of the notion of 'scientific (causal) law' to the social sciences. The defender could reply that there is no problem with this extension, since in fact it went in the opposite direction, from human behaviour to the behaviour of the natural world.

An even stronger case for the claim that the meaning of 'cause' permits legitimate application to the social world can be made by considering an alternative account of the origin of the notion of 'law'. Much can be learned on the question of the applicability of 'cause' to the social sciences by examining the earliest origins of scientific thinking and views of nature and society in Greek thought. Hans Kelsen, one of the twentieth century's leading legal theorists, wrote extensively on the origins of the notions of 'causality' and 'law'. In Kelsen's view, the traditional explanation of the origin of 'cause', alluded to above, is flawed, since early thinkers must have had a pre-existing conception of 'law' in order to connect will and behaviour. He argues that the predominant conception was not individual but was social, or perhaps religious. Kelsen argues that the origin of the scientific ideas of 'causality' and 'law of nature' are derived from social and mythico-religious sources. Because of the centrality of the origins of the concept of 'cause' for evaluating the criticism that reasons do not fit into the Western scientific concept of 'cause', it is worth offering a concise summary of Kelsen's argument.

Kelsen's argument that 'cause' in natural science derives from social discourse

Kelsen argues that people did not always possess the now-common Western notion of the causal interconnectedness of nature and that the Western scientific notion of 'cause', along with the origins of Western science, arose in ancient Greek thought. He shows that the Greeks, prior to their investigations into nature, had a much more fully developed notion of social order than of natural order. Drawing on evidence from Greek views of myth, religion, language, and especially cosmology and philosophy of nature, Kelsen argues that the Greeks developed the scientific notion of 'nature' by extending the already deeply accepted notion of 'social interconnection', and specifically the juridical notion of 'law', to account for natural observations. In order to see the power of Kelsen's argument, this section attempts to lay out his analysis of the original meaning of 'cause' in the earliest Western scientific speculation, as advanced in his 1939 paper 'The Emergence of the Causal Law from the Principle of Retribution' (Kelsen 1973).

Naturalism in contemporary philosophy of science attempts to bring some of the intellectual prestige and success of the natural sciences to the social sciences. It has been noted above that in the modern world many thinkers attempt to import the very precise methods for understanding the physical world into the social world. Kelsen sees the Greeks as doing the opposite, namely, as attempting to import the much more developed methods for understanding the social world into the natural world, as the latter had only recently come under systematic investigation.[21]

One key element of Kelsen's argument is that the social order was fundamental to Greeks, and an event that might violate that order would subsequently require a retributive punishment-event. This can be seen in virtually all of the pre-Socratic natural philosophers. Kelsen cites, e.g., Thales, Empedocles, Anaximander, Anaximines, Heraclitus, Parmenides. Kelsen's argument is also grounded in the Greek focus on harmony and moderation. In the social order, evil actions create an imbalance; harmony and justice are restored by punishment of the evil actor. In the natural order, when the weather turns colder and colder in autumn and winter, an imbalance arises. Balance is then restored by the temperature becoming warmer and warmer in spring and summer. The cycle repeats, as recurring imbalances are restored to balance.

The centrality of the concept of 'balance' in the Greek understanding of the demands of justice is evident in the doctrine that the severity of punishment for a crime should be commensurate with the severity of the crime itself. There is a similar Greek idea of the 'balance' between 'cause' and 'effect', which entails that a greater effect cannot be produced by a lesser cause. Perhaps the two doctrines of balance were connected in the early development of Greek thought. If so, the parallel suggests the likelihood that the Greek understanding of 'judicial balance' led to the understanding of the 'proportionality of cause and effect' because, as Kelsen stresses, contrary to the usual assumption, in historical development the former long-preceded the latter.

Balance can be seen in Anaximander's explanation of all things in balance as arising from the Unlimited, which is the source of things in the world, and is that to which they return when they have lost their particular nature (like fire mixed with water). Kelsen cites Capelle and Jaeger to support this line of interpretation. Capelle holds that one finds in Anaximander the first instance of a genuinely causal law, specifically, in 'the concept of a legality immanent in and governing all that happens, i.e., the entire world-process; in short, the idea of a world-law' (Capelle 1938: 97, cited by Kelsen 1973: 172). Kelsen adds that it remains, at bottom, the law of retribution. Jaeger (1939: 157ff.) notes that Anaximander holds that there is a certain 'legal status' of the 'injustice' that things do to one another as they come into being and pass out of being, which is that they 'pay penalty and retribution to each another'. Jaeger holds that this view is modelled 'on the Greek city-states' legal ideal of the "polis", which was binding on each individual' (Kelsen 1973: 204 n. 27).

The centrality of the concept of 'balance' can be seen by noting that the meaning of the term 'ἀρχέ' includes an idea of 'likeness' that 'appears as that of *balance*, which – so far as it means *justice* – is the specific function of retribution, which weighs out punishment for guilt, reward for merit, as if in the scales, and holds the balance between them' (Kelsen 1973: 171). Anaximander starts from an infinite substance that, from itself, produces opposites, like wet and dry, hot and cold. These opposites do battle with one another and any predominance by one over the other constitutes an injustice. The restoration of balance from the move toward the domination of one of them is 'a kind of retributive justice' (Kelsen 1973: 171). The same can be said for the Greek understanding of medicine. Kelsen (1973: 204, n. 24) cites Alcamaeon of Croton, who holds that health is the equal balance of moist and dry, hot and cold, sweet and sour, etc. Kelsen also alludes to the Eleatics' denial of the reality of change. Change is a problem, since from a position of stability, harmony or balance any alteration would create an imbalance. The Greek belief in balance was so powerful that it served, at least in part, as grounds for claiming that all change is illusory.

Furthermore, 'ἀρχέ' means both 'rule' and 'beginning' because the notion of the 'beginning of the world' is connected to that of 'a creator of the world who rules it'. When the earliest pre-Socratics, like Thales, Anaximander and Anaximenes, seek an ultimate foundation or principle (water, the Unlimited and air, respectively) that permits a unified explanation of the observable world, they are, Kelsen argues, seeking something that 'rules the world, as a monarch does' (Kelsen 1973: 168). Anaximines says, 'As our soul, being air, holds us together and *controls* ... us, so does wind [or breath] and air enclose the whole world.' Kelsen adds, 'When Anaximines takes the soul to be an air-like thing, it should be noted that he holds that "air is a god," i.e., endowed, no doubt, with reason and will. In this sense air "controls" the world as a basic principle' (1973: 168).

Heraclitus reverses the natural picture of Anaximander, but follows him in viewing nature through the lens of social interaction. Heraclitus says that '[w]ar is common and right is strife and that all things happen by strife and necessity'. Thus the elements of the natural world achieve balance as a result of the war they wage upon one another. (With this relationship between social reality and nature one may interpret Heraclitus' well-known claim, 'War is the father of all and king of all.') Kelsen argues that the common use of 'law' in both natural and social contexts (juridical and natural law) stems from the origins of the concept in ancient Greece, where the parallel may have been drawn between the regularities of the natural order and the regularities of Greek social order. This conception of the natural world informs Heraclitus' view of astronomy, according to which the sun follows its path in obedience to a moral law of the gods. Heraclitus says, 'The sun will not overstep his measures, but if he does, Erinyes [demons of revenge], handmaidens of Dike [goddess of justice] will find him out' (cited by Kelsen 1973: 173). Kelsen observes that the significance of this fragment of Heraclitus for understanding the origin of Western science is that 'the inviolability of the causal law whereby the sun keeps his path is the coercion of Dike, the binding force of the legal norm, a normative necessity' (Kelsen 1973: 174). It is not inconceivable that the sun overstep his measures. Retribution is what prevents it from occurring in fact.

In addition to the evidence Kelsen adduces from Greek philosophy of nature and etymology, he presents a number of examples from Greek mythology and usage to support his account of the social origin of the notion of 'causal law'. With respect to the evidence from mythology, Kelsen observes that when one god acts, that god is held in balance by the vigilance of other gods, which mirrors the human relationships in a social structure in which actions outside of circumscribed boundaries produce a retributive reaction. Likewise, Empedocles' doctrine of the transmigration of souls arises from the doctrine of retribution. The four elements of nature themselves punish the transgressor of good. Nature exacts the retribution (Kelsen 1973: 176). Parmenides asserts that the law-like necessary connections in the cosmos are 'the absolute binding-power of a divine legal norm, and … this norm – the law of nature, as the law of eternal being – is retribution' (Kelsen 1973: 175).

While the earliest Greek philosopher who is regarded as scientific is Anaximander, the earliest ones who are seen as 'purely scientific' were Leucippus and Democritus, the latter of whom first separated efficient causation from the principle of retribution, teleology and final causation. Thus in the writings of Leucippus and Democritus the modern notion of 'causation' first appears. 'So long as the world-order is construed on the analogy of the social order, as the expression of a more or less personally conceived, rational and thus purposively functioning will, the law of all that happens must have the character of a norm which, on the analogy of legal ordinance, the basic social norm, guarantees the normal order of things by means of sanctions; in short, the world law must be a law of retribution' (Kelsen 1973: 180).

Aristotle (1984: 1218; *De Gen Anim* 789b2) says that Democritus, ignoring the final cause, refers all the operations of nature to necessity. For Democritus 'causation', emanating from the interactions of atoms with one another, is necessary and purely mechanical. Still, Democritus, heavily influenced by Heraclitus (who holds, as noted, that all elements are in strife and at war with one another) sees atoms in conflict with one another. Furthermore, although Democritus treats the physical world as purely mechanical, it nevertheless follows closely the model of the retribution principle according to which 'an action is linked with its specific reaction, guilt with punishment and merit with reward' (Kelsen 1973: 183). Democritus used the concept of 'causation' in remarks such as, 'They say that nothing happens by chance, but that everything which we ascribe to chance or spontaneity has some definite cause.' The term he uses for 'cause', Kelsen notes, is 'τι αἷτον'. Herotodus uses the term in a similar way. Kelsen points out that a century earlier this term, as used by Pindar or Aeschylus, meant 'guilt'. While Homer does not use the term, he uses the related term 'αἴτιος' to mean 'guilty'. Kelsen strengthens his case by noting that at the same time Leucippus was developing the idea of 'causation' as detached from myths of retribution in natural philosophy, Progatoras was doing much the same in the realm of social philosophy, arguing that punishment by governmental agency was justified on grounds of prevention of future violations rather than on grounds of retribution (Kelsen 1973: 181–2).

Consequently, Kelsen's analysis of the origins of the notion of 'causation', show that it arose with the Western development of science and was an extension of the pre-existing understanding of social relationships among people. It is not a form of understanding inherent in all rational thought. It arose at a particular time and began to fade away in the Middle Ages until modern scientific thought pulled away once again from religious conceptions of nature.[22] It did not initially arise as a part of any philosophy-of-science-oriented understanding of the natural world, but rather as part of the Greek understanding of the social world.

Conclusions

This chapter has argued that some reasons are causes and that to construe reasons as causes does not violate the core meaning of 'cause' in the Western scientific context. One answer to the question why the US invaded Iraq is that Saddam Hussein did not allow full UN inspections from 1998 onwards because he wished to avoid appearing weak to Iraqis and other Arabs in the Middle East. Reflectivists and HT critics generally deny that this sort of reason can properly serve as a cause of the US invasion. One might also wonder how and why Britain chose to work cooperatively with the US on the invasion. Could it be that the two states or two leaders had a 'common interest' in ridding the region of a bellicose leader who was believed to possess chemical and biological weapons? Perhaps the UK, but not France, Germany or Russia, regarded the demonstrations of US hegemony as beneficial, or at least not harmful, for its national interest. This chapter has argued that factors like 'common interests' or 'perceptions of US hegemony' are proper candidates for causal forces in IR.

The US and UK, of course, do not always work cooperatively on security issues so a generalisation to that effect would have to be probabilistic (given coordination failures concerning Suez, Skybolt, the Falklands, etc.). So such explanations may be probabilistic rather than deterministic. In support of this position, this chapter has argued that there is a meaningful notion of 'probabilistic causation'. So, factors like 'common security interests' and 'hegemonic structure of international distribution of capabilities' may be regarded as legitimate candidates for causal influence, even though their associations with the states' record on reaching accord or acting cooperatively are only probabilistic. Only through further enquiry may one answer the question of which among the plausible factors were in fact the reasons or causes of US and UK behaviour.

This chapter has also distinguished the origins of meaning of, proper interpretations of, and evidence for claims about 'causality'. While the origin is derived from social relationships in the works of the first natural-science-oriented writers in the West, the interpretations are along the lines of 'relative frequency' in some cases and 'degree-of-belief' in others, and the evidence is drawn from observations about repeated instances, though not necessarily universal invariance, between one type of event and another. Those who believe that the social sciences should seek to identify underlying causal mechanisms must hesitate to accept Little's claim that such a quest precludes the search for associations. It is only by discovering those regularities that many causal mechanisms will become apparent, or even be suggested as potential mechanisms.

Human rationality and desire form a link between 'cause' and 'effect' for many social actions. The desire of President W to get re-elected is a reason for, and can be construed as a cause of, the decision to invade state X. This is true particularly in the case of first-image explanation and prediction. One can answer important questions about large samples and how they play a role in social behaviour by understanding probability statements along the lines described on pp. 93–100, and by seeing social causal relations to hold between types of events. The investigator or policy-maker may adopt propositions with different degrees of corroboration and different levels of confidence. The former are understood in terms of 'relative frequency' or 'degree of entailment', the latter in terms of 'subjective probability' or 'degrees of belief'.

Theoretical explanations enhance the investigator's confidence in the truth of a causal claim beyond what correlational analysis permits. This study endorses the notion that there is a deep structure of the world, both in terms of its physical and social character. But it holds, as all Duhemian conventionalism does, that empirical scientific laws, like other (inter-subjective) empirical scientific claims (e.g., observation reports), are never indubitable. Any acceptable causal claim must support counterfactuals, which are widely understood by philosophers to be interpretable along the lines of Kripke-style semantics (Kripke 1963). And it is possible to operationalise such statements, e.g., about Britain's arms race with Germany (in part) causing war or Britain's democratic structure (in part) causing its war-avoidance with Iceland in the Cod Conflict in 1975–6. The generalisations may be evaluated by looking at other examples of arms races and the subsequent relations between the competing states or by looking at other pairs of mature democracies interacting with each other.

This study defends the claim that conventionalism is consistent with an account of 'causation', in particular, a form of Little's causal realism (Little 1993a: 197–8). But on the account defended here causes are real, though causal claims are relativised to a body of evidence. As the evidence changes, the understanding of causal powers changes, because the latter are based on claims about association and probability. As new observations are added, the probabilities change. If $p(A/B\&E) \neq p(A/E)$, then B is causally relevant to A on evidence E. But this conclusion is fallible. Just because B is taken as causally relevant to A on a given set of evidence, it may be that new evidence will undercut that belief. For example, further study may produce a new evidence set, E'. When E' replaces E, it may be that $p(A/B\&E') = p(A/E')$, in which case the claim that B is causally relevant to A would be rejected. Just as new evidence may establish a causal connection, it may also disestablish one.[23]

The long and intense philosophical debate over the meaning of 'causality' has led to a popular view of the term as intrinsically deterministic and thus as incapable of being intelligibly applied to the social world. But the natural sciences have been increasingly interpreted as non-deterministic. And Kelsen's study of the origins of the concept of 'cause' in Western science lends support to the idea that its application to the social world is not an unacceptable stretch from its core meaning as a scientific/deterministic concept, since historically the notion was stretched in the other direction – from its origins grounded in application to the social world to a derivative application to the natural world. As Jervis points out, theoretical explanations must remain part of the basis for cause-and-effect claims, which supports the notion that theorists must keep a vigilant eye on the evidence on which hypotheses are accepted. A conventionalist account of theories allows one to make sense of 'causality', to make sense of the 'truth' of laws of science and to avoid the charge of arbitrariness in our choices among theories.

5 Prediction, theory and policy-making

TESMAN: What, the future? But good heavens, we know nothing of the future.
LÖVBORG: Yes, but there is a thing or two to be said about it all the same.

Ibsen, *Hedda Gabler*, Act II, tr. Edmund Gosse and William Archer

'It's exactly the opposite of what we intended,' a senior Administration official said this week. 'In retrospect, perhaps it was predictable. But very little of what North Korea does is ever predictable.'

David E. Sanger (*New York Times* 28 December 2002: 11)

Formulating security policy regarding intervention against and rebuilding Iraq will require answering questions like whether toppling Saddam Hussein's government will reduce the incidence of terrorism against the West, whether a long post-invasion occupation will lead to a greater chance of a stable, pro-democratic Iraq, etc. The answers can only be used as the basis for policy-making if predictions are derived from the causal claims (as Chapter 1 argued) and are only possible if generalisations are developed (discussed in Chapter 3) which have a causal character (as Chapter 4 argued).

IR theorists quite often draw predictions from the regularities they propose. All of the authors discussed in Chapters 1, 3 and 4, who examine the 2003 war in Iraq offer predictions. Kenneth Pollack predicts correctly that the worst problems 'for the United States are likely to stem not from the invasion but from the aftermath' (2002). His prediction is grounded in his generalisation regarding the difficulties of nation-building, which in turn is based on observations from past cases. Dobbins et al. (2003: 205) also offer predictions, based on three lessons that embody laws and regularities:

1 The co-opting of existing institutions can facilitate democratic transformation but the results may sometimes be less thorough-going than starting anew.
2 Elections are an important benchmark of democracy. Held too early at the national level, they can strengthen the extremist and rejectionist forces rather than promote further transformation.
3 Imposed justice can contribute to transformation.

The authors argue that these three principles may be applied to policy toward Iraq. Thus, they predict that transition will probably go more smoothly and successfully if justice can be imposed by bringing to trial senior figures who were responsible for past crimes. And they predict that the transition will go more smoothly if elections are held, but not too quickly.

Byman (2003) argues that most of the barriers he cites are connected to problems of security, which he regards as potentially soluble by proper US policy choices. Thus he offers predictions, but they are, like those of Dobbins et al., conditional predictions that are contingent on US policy decisions. For example, Byman says (2003: 74–5), if the US commits sufficient resources to rebuilding Iraq's security infrastructure, Turkey and Iran will not succeed in de-stabilising a new Iraqi democracy. Noah Feldman (2003) also offers many predictions about Iraq. He begins with two general propositions. The first is that the will of the majority is manifest in legislation in democracies. The second is, 'A great majority of Muslims do not believe that Islamic practices should be enforced by the state' (2003: 228). From these he derives the prediction that a democratic regime in a Muslim society would not legislate strict, religiously inspired, limitations on individual freedoms (2003: 228). Hollis offers a number of predictions, one of which is that US control of Iraqi oil fields will lead to a failed transition to democracy (Hollis 2003: 31). Brooks (2002), Metz (2003–4) and Barton and Crocker (2003) all similarly offer predictions derived from their causal analyses, noted in Chapter 4.

Is prediction of the sort needed here even a possibility in the social sciences? While naturalists have said it is (see Chapter 2), there are many critics who doubt it. Some emphasise the notion of 'meaning' in social science theory and the hermeneutic tradition (HT) that builds upon it, some focus on the complexity of social world or social sciences, and others focus on the non-linearities of social behaviour.

Many authors have supported the idea of the predictiveness of social science theory. As long as people have been studying IR, there have been authors who were confident of their ability to isolate the key causal factors that determined outcomes and to develop predictive theories. Twenty-five centuries ago Sun Tzu boldly asserted, 'If you say which ruler [possess which characteristics] … I will be able to forecast which side will be victorious and which defeated' (Sun Tzu 1971: I 11–14).

Much of early twentieth-century philosophy of science was influenced by the h-d method, which envisioned testing of theories by deriving hypotheses, creating conditions under which a predicted event would occur in the form of experiments, and then observing whether the specified outcome occurs. If so, the evidence counted in favour of the theory; and, if not, against it. This traditional method casts prediction in a central role. Such a view was supported initially by logical positivism, which attained hegemony for a time, focused on the paradigmatic status of the natural sciences and stressed the common methodological character of the natural and social sciences, i.e., the 'unity of science'. Logical positivism soon faced criticism from so-called logical empiricism, such as the inability to account for 'prediction', even though logical positivists endorsed theory-based prediction.[1] If science, as the positivists argued, is based on the laws of logic and the past sensory experiences of individuals, then there is no guarantee, or even probability, about what will happen in the future. Scientific laws secure prediction (deterministic or probabilistic) and to logical empiricists like Hans Reichenbach it was unclear how logical positivism, e.g., in the form of Carnap's *Aufbau* (1967), could account for this. Reichenbach said, reviewing that work, 'Indeed, there is no scientific law which does not involve a prediction about the occurrence of future events; for it is of the very essence of a scientific law to assure us that under certain given conditions, certain phenomena will occur (Reichenbach 1936: 152. See also Clendinnen 1979: 100–14).

Many quantitative modellers, rational-choice theorists, and others retain something of Sun Tzu's and logical positivists' confidence in predictiveness. Prediction-optimism was buoyed in some quarters of IR by the rise of behaviourism and statistical modelling. Indeed, most authors who have had a major impact on the field in the past half-century have endorsed predictive theories of IR. Ray and Russett (1996) have called attention to the fact that many predictions of major events have been accurate. An ambitious recent example is Bueno de Mesquita's *Predicting Politics* (2002). Is there some firm foundation for a belief that accurate prediction in IR and the social sciences is possible and not merely the result of luck?

In recent years, however, a number of IR theorists have expressed scepticism. Ashley (1986), Cox (1987), Walker (1993) and, most recently, Patomäki (2002) have emphasised the interpretive nature of social theory in disparaging prediction in IR. Wendt's constructivist approach emphasises scientific-style explanation, but not prediction. Robert Jervis has argued that causal mechanisms should be identified instead of superficial generalisations from observed events, and that such generalisations offer facile routes to predictiveness. The distinction Jervis draws is used by philosophers of science to criticise theories that claim predictiveness. Jervis (1976: 415) himself disparages but does not dismiss predictiveness in IR. Power-transition theorists like Gilpin (1981) and Doran (1999) have also criticised predictiveness of IR theory, in Doran's case based on the non-linear nature of the transition phases.

Philosophers of science typically regard prediction as a standard part of natural-science theory. The principle of symmetry between explanation and prediction, which clearly exhibits the legitimacy that prediction has enjoyed in the account of the natural sciences, had considerable popularity for a good deal of the twentieth century (see pp. 39–42). Historically critics have questioned the appropriate form that prediction should take but not whether it has a place in the natural sciences. Even after much critical scrutiny and rejection of the symmetry principle, predictiveness remained an essential part of the enterprise of science.

In the past several decades the post-positivist meta-theoretical literature has shown an increase in opposition to social prediction. Constructivists, reflectivists and critical theorists in IR generally reject any notion of 'predictiveness'. They either do so implicitly or explicitly, like Cox, who says, flatly, 'It is impossible to predict the future' (1987: 139, see also p. 393).

The implicit rejection of 'prediction' is widespread among constructivists and reflectivists. Wendt and Walker are clear examples. Walker has little to say about 'prediction'. While he offers a sustained critique of naturalism and the empiricist (though not empirical) approach to the social sciences, he focuses on the logic of explanation and the presuppositions of the dominant forms of theory rather than questions connected to 'prediction'. Wendt (1999) lays out his extensive meta-theory barely ever mentioning 'prediction'.

Ashley inquires about the possibility of predictive theory in the social sciences. He asks, 'how can there be a naturalistic social science, one that produces objective knowledge capable of calculating and predicting social outcomes, given that human action is necessarily "subjective" in character?' (1986: 283). Ashley notes that the answer Weber (1974) gives is that 'we abstract and regard as objectively given an agent's substantively empty logic of technical reason', which justifies the assumption that 'society will appear to the individual agent as a subjectless set of external constraints, a meaningless second nature … [Thus] knowledge of an agent's pregiven ends and meaningless social constraints, meaningful and "rational" subjective relations become calculable, predictable, and susceptible to causal accounts' (Ashley 1986: 283).

The most recent reflectivist work on IR meta-theory, that of Patomäki (2002), rejects prediction. Patomäki (2002: 157) says that 'qualitative changes and emergence are possible, but predictions are not' (and reiterates this, 2002: 191). Yet on the next page Patomäki (2002: 158) begins a one-page discussion of the performance of political action. He mentions 'prediction' on a few occasions simply to note that it does not work in social science theory. He says (2002: 135) that 'mathematical or statistical models … are highly unlikely to be able to predict the future'. He adds that rational choice models also fail (2002: 168–73).[2]

The trend towards rejecting or downgrading 'prediction' is reinforced by various developments in current philosophy of social science.[3] This chapter considers the difficulties for the contemporary analysis of policy formation that arise from such anti-positivist and reflectivist thought – since theory-based prediction was a hallmark of traditional positivism. The chapter examines three different sources of prediction-scepticism in the philosophy of social science, drawn from the indeterminacy of social theory in the first section (Max Weber, Jürgen Habermas, James Bohman and Bernstein et al), the lack of governing regularities in the social sciences in the second section (Nancy Cartwright and Daniel Little) and the effects of non-linearities in the third section (Charles Doran), the mistaken analogy with the natural sciences in the fourth section, and offers a sketch of how a conventionalist approach solves one of the central problems in the final section.

It will be argued that all three sorts of anti-predictive argument are flawed and that the latter two presuppose an indefensibly narrow notion of 'prediction'. Since all policy-formulation is future-directed, as it is an attempt to influence what will happen in some time to come – near or distant – some connection must obtain between present actions and future outcomes. The need to predict those future outcomes, given initial conditions (and intervening actions), is rarely addressed by anti-positivist IR theorists, and still less by anti-positivist philosophers of social science on whose work IR theorists have relied.

The hermeneutic circle

The hermeneutic tradition and Weber's account of social theory

One critique of 'prediction' in IR comes from a view of the character of social science that fundamentally opposes the naturalist perspective and that instead sees the purpose of social enquiry as the extraction of meaning from the events of the social world. The hermeneutic approach to IR, introduced above (pp. 24, 44–51) treats social analysis in a way that does not acknowledge a place for prediction. Are hermeneuticists correct in disparaging social prediction?

Hermeneutics originates in the attempts to understand the meaning of holy scripture. Such attempts have a unique set of problems associated with interpretation. The methods developed were extended in order to a devise a method for understanding texts in general, and later, especially by Wilhelm Dilthey (1996), to works of art, human life and to history and human action. The effort has been taken up by Heidegger (1965) and Gadamer (1976) in the twentieth century and applied by many authors to the social and political worlds.

The critique offered by the HT covers all schools associated with positivism by arguing that the scientific approach to the social world ignores the centrality of the notion of 'meaning', in particular the meaning of actions and the ways in which events or behaviours are understood by the participants. In order to understand an agent's action, that action must be located within a particular conventionally defined symbolic performance, which must, itself, be located within the more general system of symbol-laden actions. The utility of the interpretive approach depends on the degree to which proponents manage to identify symbolic social structures in human behaviour.

HT authors hold the process of interpretation, whether of the Bible, works of art or political actions, to be circular in the following way: if a system of interpretation is necessary before an action can even be identified properly as the action it is, then the classification of the action must be relative to that particular framework of meaning. One must then interpret in order to ascertain the system of meaning. Yet one must have a system of meaning in order to interpret. Similarly, with a text, one must understand the whole of a work to understand the meaning of specific sentences, but one must understand the specific sentences of the work in order to see the meaning of the whole. The hermeneutic circle, they argue, is inescapable.

HT authors oppose naturalism in part because the social sciences are, in their view, incapable of natural-science-like generalisations. Authors, like Alasdair MacIntyre (1973), deny that universal or cross-cultural generalisations are possible, since action need not in all cultures seek to fulfil material needs and other identifiable goals. One must understand the symbolic forms of the culture before interpreting the political or economic behaviours. Symbolic forms constitute the terms and character of social action. Because each social institution in a particular society is dependent on the specific characteristics and symbols of that society, there is no universal or abstract notion of an 'institution' that can even be meaningfully applied from one culture to another.

HT writers acknowledge the seminal contribution of Max Weber, who argued that social science theory must be adequate on the level of both science and meaning. Parsons (1937: ch. 8) interprets Weber as harking back to the tradition of German idealism in his reaction against the growth of positivism during his lifetime.[4] Social science generalisations must confront the notion of 'meaning' if they are to be more than mere statements of statistical association. A focus on meaning implies that the statements should show internal connections between rules that govern action. Generalisations must also be supported by empirical evidence to show that the connections they assert or the motivations they hypothesise obtain in the real world. Each sort of generalisation requires the other to be fully satisfied. They require meaningful causal claims (or reconstructions of event-types) as well as probabilistic evidence.

Hermeneutics maintains that action must always be understood 'from within', which means that the investigator needs to know what the agent's intentions are in performing an action and what the rules and conventions are that govern that action. HT theorists require that interpretation proceed in two ways: to understand the action of an individual, it must be interpreted in light of the intentions of the agent, and to understand social practices, it is necessary to understand the significances that practitioners attach to them, as the latter in part constitute those practices.

HT authors object to behaviouralism by pointing out that the same action, with one and the same causal path, may have different meanings and may be intended to convey different messages when performed in different contexts (that are not causally connected). They conclude that the differences in context cannot account for the differences in meaning. The actions of a chess player are only comprehensible by reference to the rules of chess (perhaps conjoined with the intentions and motives of the player), which, according to the HT, are not causally connected to the actions performed. Rules alone do not account for all aspects of two agents' interaction over a long period of time. The choice among various possible courses of action, some of which are only subtly different and all of which fit within the rules, will have a bearing on meaning and behaviour. Thus Pierre Bourdieu (1977) says that, in order to understand two agents dealing with one another over a long period of time, the analogy of two boxers in a ring is more illuminating than simply that of a set of rules. The rules create the framework in which the boxers interact with one another. But the vulnerabilities and motives of the boxers also contribute to the resultant behaviour.

In the HT, social processes are brought about by agents, who have norms, values, beliefs about the appropriate rules of behaviour and understandings of their own power and limitations. The goal of social enquiry is explanatory understanding, which is to support an hypothesis about the beliefs, values and purposes that bring the agent to act as he or she does. The goal is to place the action in a broader context of meaning. The process involves considering alternative plausible states of the agent's mind and then applying the available evidence to determine which hypothesised state of mind holds up best against it.

For Weber, agents act in a context of shared social meanings and rules but each actor proceeds according to his or her individual goals and motives. Weber is individualist in this respect, in contrast to other hermeneutic theorists, like Durkheim (1972), who tend to stress social structure. For Weber collective action results primarily from individual decisions rather than from social structure. According to Weber, all social actions are guided by the actor's intentions and are oriented toward other agents. The phenomena one observes are partially constituted by the intentions of actors. While one can formulate correct generalisations and abstract propositions, they are of little value because the abstract statements do not help in the understanding of phenomena in a particular culture or setting.

Bohman's hermeneutic account of the social sciences

James Bohman (1993) offers an account of social science that emphasises the centrality of 'interpretation' and the inescapability of the hermeneutic circle (see Habermas 1971, 1973, 1984). Bohman's account is thus squarely in the HT. Bohman argues throughout that all attempts to escape the hermeneutic circle are not only futile but unnecessary. Interpretation is always a part of the description of a social action. Bohman especially aligns his extension of hermeneutic thinking with that proffered by Habermas. Bohman (1993: 111) contends that because the circle is inescapable, research programmes that attempt to avoid interpretive circularity (like ethnomethodology and rational choice) must fail, ultimately replacing one type of circularity with another.

Some HT writers offer extreme versions that leave the social sciences with a range of undesirable consequences, especially severe limitations on what they regard as knowable. Bohman (1993: e.g., 124–6) sees himself as rescuing social theory from the excesses of previous interpretivist and hermeneuticist scepticism. Bohman seeks to show how major social theorists (e.g., Durkheim, Weber, Marx, Parsons, Taylor and Winch) have not succeeded in dealing with various types of indeterminacy, which arise from sources such as interpretations, causes, rules, macrostructures and criticism. Bohman's theory draws primarily on the first, interpretive indeterminacy.[5]

Bohman holds that most earlier HT theorists erred regarding the limitations on knowledge and the use of evidence in the social sciences. Bohman's work is especially worthy of examination because, in comparison with others' work in the HT, his is moderate and less epistemically limited, and he regards the HT as offering an optimistic account of knowledge. Bohman says that the new philosophy of social science has accepted indeterminacy and made 'it manageable within empirically adequate and verifiable explanations' (1993: 232). The scepticism that many proponents of hermeneutic interpretation come away with is, according to Bohman, based on a one-sided analysis of one sort of interpretation: 'holistic-contextual' interpretation, which ignores 'rational-comparative' interpretation. This study seeks to avoid attacking a 'straw person' by selecting the less radical version of Bohman. Examination instead of the more radical sceptics in the HT would leave open the possibility that others in the HT offer a more expansive theory of knowledge that might be held to be capable of supporting prediction in the social sciences.

Howsoever optimistic Bohman is about a variety of knowledge claims, he, nevertheless, does not find justification or legitimacy for prediction in the social sciences. Bohman clearly sees avoidance of scepticism as a central goal of his account of the social sciences. He says, 'This book is profoundly anti-skeptical' (1993: 14). Bohman flatly rejects that the social sciences are predictive. In the first chapter of his book he states that the social sciences 'fail to yield unique predictions' arising from determinate laws like those in the natural sciences' (1993: 13). After rejecting positivism and explanation–prediction symmetry in that chapter, Bohman has essentially finished his discussion of 'prediction'. Clearly anti-scepticism is a conclusion to which he is fully committed. But his book is sceptical about prediction. This is where his conception of the social sciences falls farthest short of the epistemic power and potential of the natural sciences.

Bohman does accept that the social sciences are explanatory in something like a scientific sense. Within its indeterminate realm the social sciences can serve up 'adequate and fruitful explanations that can fulfill a variety of purposes' (1993: 13). Moreover, causal explanations are possible, if causal mechanisms can be discovered. According to Bohman, the discovery of causal mechanisms is possible in the social sciences, but the formation of general laws is not (1993: 111). Thus Bohman goes on to say that even if causal mechanisms are detectable in the social sciences, they are not general enough to yield predictions. He says, '[a]dequate causal explanations of actions are still possible if we can discover the appropriate mechanisms with specific empirical scope; such mechanisms, however, are not general enough to permit predictions and may themselves be altered by agents who become aware of them' (Bohman 1993: 233).

According to Bohman, predictive success requires determinacy (1993: 7). Since the rest of the book (subtitled: 'problems of indeterminacy') argues for indeterminacy in the social sciences, it is clear that he holds prediction in the social sciences to be impossible. Bohman does not seem to allow for probabilistic or negative forecasts (ruling out that certain events will occur, like a near-term French invasion of China, which are considered on pp. 142 ff.). Either of these might justify predictions of a limited nature. Bohman is explicit and unqualified in his rejection of rationally justifiable prediction in the social sciences.

Bohman's account of 'rationality' and 'evidence'

In rescuing the HT from its more extreme forms of anti-naturalism, Bohman develops notions of 'evidence', 'justification' and 'rationality' that allow for a useful (but not naturalistic) notion of 'explanation' in the social sciences. Bohman thinks other analysts in the HT place excessive limits on social-scientific knowledge because, while they correctly see that knowledge is not objective in that it is conditioned, they fail to see the difference between limiting and enabling conditions (1993: 121–4). The former are set and fixed, while the latter are variable and may even change as a result of conscious design.

Contrary to many in the HT, Bohman holds that there are intersubjectively valid ways of evaluating competing interpretive claims (1993: 107). 'When faced with competing claims many now doubt that there is only one way to settle interpretive disputes rationally or that there is any criterion to decide success or even comparative superiority' (1993: 113). Still, both contextual and rational interpretation have 'correctness' as a regulative ideal (1993: 142–3). The former places utterances in a larger whole to make them intelligible, while the latter evaluates the differences between different points of view.

Contextualist scepticism argues that interpretation is universal, i.e., all understanding is interpretation, and that its presuppositions cannot be specified because interpretation takes place against the background of all of one's beliefs (Bohman 1993: 142–3) The assertion that 'x is the correct interpretation' is itself an interpretive claim. Interpretation is therefore 'indeterminate, perspectival and circular'. But Bohman (1993: 112) asserts that there can, nevertheless, be knowledge based on evidence, and he disagrees with those who draw more sceptical conclusions from the inescapability of hermeneutics.[6]

One important question is whether it is possible on the basis of rational considerations to decide the merits of competing interpretations of events. Is it ultimately subjective, is it ultimately arbitrary, or are there intersubjectively valid methods for deciding upon an interpretation? Habermas holds that all interpretation serves a single function (and that 'all interpretation is evaluation'). And significantly, Bohman disagrees with Habermas on this point, since Bohman recognises multiple functions for interpretation. In so doing, Bohman paves the way for a more intersubjectively valid notion of 'rational decidability between competing interpretations'.

Bohman stresses that the background sets of beliefs, which are necessary for the possibility of interpretation, are not limitations on knowledge but rather are 'enabling conditions', which are shared, public background constraints. These constraints 'are not strong enough … to make it impossible to decide normatively between interpretations on the basis of evidence' (1993: 125). Bohman adheres firmly to an intersubjectively valid notion of 'knowledge', which is produced through the process of theory choice and comparative evaluation of interpretations. For Bohman, knowledge is 'fallible and revisable'; further evidence and analysis may always produce new conclusions about the comparative evaluation. But the position Bohman develops allows that 'better and worse interpretations can be established' (1993: 125), since the warrants for knowledge claims can be debated and decided publicly. This provides Bohman with a much less limited theory of knowledge than – and sharply separates him from – others in the HT.

A basis for prediction within Bohman's theory and the hermeneutic tradition

While Bohman himself rejects prediction, his argument against it appears flawed and leaves the way open for a notion of 'prediction' within his indeterminacy-based account of social-scientific knowledge. His rejection of the symmetry between explanation and prediction appears to be in part a result of the very different types of analytical assessment they receive; it seems less a conclusion of systematic analysis than an artefact of unequal treatment. Bohman essentially dismisses the question of 'prediction'. He rejects it with little investigation (as noted, after his discussion of positivism in his first chapter, 'prediction' disappears). But he goes to great lengths to salvage a notion of 'explanation' from HT criticisms.

Bohman's ideas of 'rationality' and 'evidence' can provide a foundation for a revised concept of 'prediction'. Bohman defends the idea of 'explanatory rigor' throughout the book but finds prediction in the interpretive world of the social sciences unacceptable (see 1993: 111 and ch. 3 passim). While a strict symmetry between explanation and prediction cannot be defended, they have far more in common than Bohman recognises. As noted above, Bohman's account includes the tenets that there is 'public evidence within the hermeneutic circle', that these evidentiary propositions allow comparative evaluation of explanatory propositions, and that, in the comparative assessment, correctness is a regulative ideal. Why are these propositions limited to those of interpretive and explanatory nature and thus why do they exclude those that express predictions? The background against which predictions are made, a background whose limitations are inescapable, is no less public and accessible than the background of interpretation and explanation.

There are reasons to question the HT approach in general and its application in IR in particular. Nevertheless, for those committed to HT, the concept of an 'indeterminate prediction' understood in Bohman's sense of 'indeterminacy' is no oxymoron; it is a cogent notion that can provide guidance for policy. Even within the hermeneutic circle, if the conditions of the circle both limit and enable, are public and accessible, and are capable of supporting a notion of 'rationality' strong enough to allow intersubjectively valid adjudication of competing interpretations, then the conditions are quite strong enough to support sufficiently a concept of 'prediction', along the lines of the definition on p. 8. According to the definition, a prediction in the natural or social sciences is a singular or general proposition, which:

1 is indexed to the future relative to the moment of its utterance;
2 may be based on imperfect evidence;

3 is based on rationally justifiable body of theory, broadly construed;
4 may be either deterministic or probabilistic;
5 the event predicted involves the sort of phenomenon that serves as a dependent variable in the particular field; and

As time passes, more experience accumulates and new interpretive claims change the enabling conditions. On this basis the predictions may be revised. Short- and medium-term predictions (which help to lend or remove support for theories) are possible. The longest-term predictions might be subject to difficulties within the hermeneutic circle because, in the long run, redefinitions of terms of discourse develop as the conditions defining the circle meld it into a new and distinct circle (though of course it is impossible to identify the moment the circle becomes a new circle).

Bohman makes brief reference in his conclusion to how agents deal with the indeterminacy of the world. He repeats his denial of social prediction, stating, '[a]dequate causal explanations of actions … are not general enough to permit predictions' (1993: 233). He then says, '[s]ocial theories are not so much instruments for … controlling future events as they are the means by which reflective agents become aware of their circumstances and how they can change them' (1993: 234). This statement does not seem to be intended to back away from prediction-scepticism. But it runs into contradiction with his view of the value social theory holds for bringing about change. How can one make desired changes if one does not have some beliefs (which are founded, at least in part, on a rationally justifiable theory) about how the causal mechanisms will in the future (perhaps with some degree of probability) lead from policy action C to desired consequences E? Such a belief clearly qualifies as a 'prediction' according to the definition on p. 8 and above.

Later on the same page Bohman (1993: 234) seems to ease his position somewhat by saying the social sciences do not give one 'determinate control' over the future, but he never talks about any lesser degree of control, e.g., probabilistic control, etc. The use of 'determinate' appears designed (though probably not consciously) to divert attention from the inconsistency between Bohman's endorsing the possibility of 'changing one's circumstances' and his denying that any sort of prediction is justifiable. By denying only 'determinate' control Bohman seems to suggest the possibility of some lesser form of control over the future. But in fact nowhere else in the book does he endorse, mention, allude to or hint at, any such lesser form.

Consider a prediction like, 'a long-term occupation of Iraq by 130,000 post-invasion US forces is more likely to create a stable, peaceful, prosperous Iraq than a rapid withdrawal of US forces'. This statement treats only probabilities and does not meet the criterion of 'determinate prediction'. Its acceptance does not imply anything that could be regarded as 'determinate control over the future'. Yet it is the sort of prediction that policy-makers ponder regularly and which they have either to accept, reject or suspend judgement on, when making decisions (such as those that the US government faced regarding the stabilisation of Iraq). This example does not conform to the notion of 'prediction' that Bohman seems to have in mind, as best as one can tell in the absence of an explicit definition.

Habermas and Bohman emphasise the centrality of 'meaning' and 'interpretation' in the formulation of social science theories. They deny that prediction is possible. But Bohman's very detailed argument leaves an opening for a significant sort of prediction in IR, namely short- and medium-term prediction. The hermeneutic circle is stable; change is gradual over time and large changes are possible only over the long term. Long-term change in the identity of the hermeneutic framework for interpretation creates problems for long-run prediction. But this leaves the possibility of short- and medium-term predictions. The conditions that Bohman defends as necessary for acceptable explanation within the hermeneutic framework do a similar job of justifying predictions and of allowing an evaluation of competing predictions as they derive from competing interpretations of intersubjectively valid empirical claims. Thus the rejection of prediction across the board is not justified by Bohman's argument and the short- and medium-term predictions that scholars, and especially practitioners of foreign policy, typically make are not undercut by Bohman's hermeneutic framework.

The argument here attributing a basis for prediction to Bohman is similar in form to the argument he offers regarding the role of criticism in ethnomethodology. Bohman notes that 'according to some of its practitioners, ethnomethodology denies the possibility of social scientific criticism' (Bohman 1993: 191). But Bohman goes on to add, 'However, each of their patterns of explanation contains premises that make it possible to use their explanations for critical purposes' (Bohman 1993: 191). It is argued here that Bohman's account offers 'patterns of explanation' that allow a parallel argument that justifies social prediction.

A further comment on hermeneutics and prediction is in order. As Chapter 2 argues, hermeneutic analysis makes use of h-d reasoning. The form of inference involved in h-d reasoning (e.g., of the sort Hopf describes in steps a–c on p. 51) is the sort of reasoning needed for predictive inference. Problems of open systems and weak probabilistic laws must be overcome in order to have any 'expectations' whatsoever. If a policy-maker accepts a theory, whether it is interpretive or not, he or she must consider which outcomes follow from which already-accepted claims. Again, it is important to stress that the notion of 'prediction' advocated here is an extremely broad and flexible one, as the definition on p. 8 and above embodies. In order to have rational justification for the policy of invasion of Iraq, one must believe that invasion will (at least probably) lead to the ousting of Saddam Hussein. The belief is based on some articulated or unarticulated theories, principles or beliefs.

The sort of h-d inference that interpretivists (like Hopf) regard as justifiable moves from known instances to unknown instances; the temporal frame of the unknown cases (whether past, present or future) neither increases nor decreases the validity or justifiability of the inference. In terms of the metaphysics and theory of knowledge that Hopf endorses, prediction remains a possibility. Interpretivists and post-positivists are, after forty years, still reacting so dramatically against logical positivism that too often they reject whatever they see as associated with it – and prediction has always been central to logical positivism. When it comes to prediction in particular, they are correct to reject symmetry with explanation. But their unrestrained opposition leads them mistakenly to reject social science prediction, even in the broadest, most fallibilistic and probabilistic sense.

Argument from absence of governing regularities

Cartwright's distinction

Robert Jervis (1976) disparages predictiveness in IR largely because he does not find the development of theory sufficient to sustain reliable prediction. He says, 'A high degree of knowledge is needed before expectations, even negative expectations, can be stated precisely, and this requirement is rarely met in foreign policy because theories of IR are poor' (Jervis 1976: 415). Jervis argues that the generalisations that are derived from experience are often inaccurate because they identify superficial connections between event-types and do not reflect the more fundamental or underlying causal connections. The distinction between observed connections and deeper causal relationships has fuelled a second sort of argument against prediction in the social sciences.[7]

The philosophical basis for this argument can be seen in the work of Cartwright, who presents a view of the natural sciences that does not fall neatly into one of the usual categories (see pp. 100, 221n.8). For example, as was shown in Chapter 3, while she agrees with scientific realists by recognising that there is a distinction between 'observable' and 'unobservable' entities and that both can properly be said to exist, she parts company with them in her distinction between 'phenomenological' and 'fundamental' laws.[8] The latter seek to explain the observations generalised in the former. Cartwright accepts the legitimacy of phenomenal laws but not fundamental laws, since, she argues, they are not generally true, even in physics. Fundamental laws are usually only true of the objects in the theory's model, while the phenomenological laws that are those that are true of real objects (Cartwright 1983: 4, 151–8).

Cartwright contrasts the use of models in science with the use of 'cause'. She points out that there is no disquiet among scientists with multiple models, which are invoked variously to highlight different aspects of the theory or different phenomena (Cartwright, 1983: 83–5). However, when one explains causally, one takes the most plausible causal mechanism, uses it consistently and does not offer alternatives simply to suit the immediate purposes at hand. Thus causal explanation, once offered, is treated as real, while models need not and should not be mistaken for 'reality'.

Cartwright agrees with many scientific realists by accepting the cogency of abductive inference but diverges from them by accepting it only in the case of causal explanation. She holds that causal explanation proceeds by citing events that lead up to the explanandum event. The description of the events often involve unobservable entities. Causal explanations usually are not committed to any particular theoretical account of the unobservable entities but only to the 'low-level' properties of such entities. Thus causal explanations are committed to phenomenological laws, but not to any fundamental laws.

While Cartwright's account has been highly influential in the philosophy of physics, much less follows from it for the social sciences. She lays out criteria for IBE, which amounts to 'inference to the most likely cause', since it is only valid in the case of causal explanation (1983: 6). But the controlled experiments that she regards as necessary for such abductive inference are not possible in most social sciences. Thus Cartwright cuts off IBE, or to most likely cause, in the social sciences. So knowledge of future states of affairs using knowledge of the present can come only from causal connections, which require some sort of 'necessary connection' between states of affairs. However, as just noted, Cartwright limits available knowledge of such connections to those derivable from experimental procedures, which are absent in the social sciences. Let us then turn to a follower of Cartwright's central distinction, Daniel Little, who uses it to construct accounts of 'causation' and 'prediction' in the social sciences.

Little's attack on prediction

Chapter 3 noted the basic elements of Little's notion of 'law' and Chapter 4 discussed Little's endorsement of 'causal realism' in the natural sciences and social sciences (1991, 1993a, 1998). Little (1993a) argues, using Cartwright's distinction, that in the social sciences there are phenomenal regularities but there are no governing regularities.[9] According to Little, causal realism entails that the social sciences should aim to discover mechanisms that are derived from agents and institutions, which in turn produce regularities. While this study endorses some elements of causal realism, it rejects other tenets of Little's position that led him to eschew 'prediction', especially the claim that the social sciences have only phenomenal and no governing regularities and thus that the associations between variables are too weak to support predictions.

Little claims that social regularities are phenomenal and thus not explanatory. As a physical example of a phenomenal regularity Little cites the manner in which glass flows like a liquid. According to Little, the flow property is neither accidental nor essential. It is not an accidental regularity, since it supports counterfactuals; thus it qualifies as law-like. But such a regularity is neither essential nor determining because it is derivative of fundamental properties – the micro-structure of the substance. In the case of the social sciences, phenomenal regularities are derivative of features of, e.g., individual agency, and do not govern or constitute the individual agent or social institution. What properly counts as an explanation in any given situation, according to most post-positivists, depends at least in part on the context in which an explanation is presented. (Chapter 4 argued that this is the case for causal explanations and Little holds that all scientific explanation is causal.) In some such contexts phenomenal properties, which may be deemed *relatively* superficial, are what one seeks.

In stating his example Little notes that glass windows are thicker at the bottom, and he says that when seeking an explanation, what is required is the micro-structure. It is not enough to give the phenomenal answer that glass flows like a liquid. Nevertheless, one might object to Little on this point by noting that *in some contexts* it would be enough to learn the phenomenal regularity that glass flows like a liquid. For example, someone who notices the uneven thickness of the glass might conclude that the glazer performed incompetently. To counter this conclusion it suffices to learn that glass flows like a liquid; the phenomenal relationship is the best one can hope for. Suppose the window consumer, alleging substandard work, brings legal action against the glazer. The explanation offered, in the form of testimony in court from an expert witness to a jury presumed ignorant of the some properties of molten glass, is that it is typical of glass windows that they are thicker at the bottom because glass flows. The phenomenal regularity is all that is required in this context.

The example could be strengthened by adding the stipulation that the case is being contested at a time (i) when the micro-physical structures of some materials have been determined and (ii) when decades of empirical examination of samples have clearly established that window glass is thicker at the bottom, but (iii) before the micro-structure of glass in particular was fully understood. In this case, the explanation from the scientific testimony in court, which was based on the phenomenal regularity, would suffice.

There are two ways one might criticise Little's claim that in social sciences, such as IR, there are phenomenal but no governing regularities. One is to point out that classical realists in IR might well disagree with Little by arguing that 'states act so as to maximize self-interest, where "interest" is defined in terms of power', would constitute a fundamental regularity. It is true in virtue of fundamental properties of the state.[10] This would be even more persuasive a claim if one adopts Cartwright's view that the claim need be true only of the objects (nation-states) in the model. Classical realists might offer other propositions that constitute phenomenal regularities, e.g., 'states in multipolar systems pursue balancing behaviour', which follow in a derivative way from the governing regularities.

A second way to attack Little's prediction-scepticism is to focus on the fact that Little bases it on the lack of governing regularities in the social sciences. One may then ask whether the sort of distinction Cartwright defends in the natural sciences is illuminating at all in the social sciences. That is, one might ask whether it makes sense to bifurcate 'governing' and 'phenomenal' regularities or whether it might be better to conceive of the relationship among social regularities as more akin to a continuum with more fundamental regularities at one end of the continuum and increasingly more derivative regularities further along the continuum. Moreover, whether a social scientist sees social laws as forming a continuum or falling neatly into two categories might depend on the particular discipline in question. One researcher's dependent variable is, of course, another researcher's independent variable.[11]

Little warns that the 'unhelpful analogy' of the social with the natural sciences produces dangers for possible philosophy of science conclusions. Yet he appears to exhibit a parallel confusion. There is a danger of uncritical analogy between the philosophy of social sciences (developed by Little) and the philosophy of natural science (developed by Cartwright). According to the latter all regularities fall into one or the other of two distinct categories and the standard for acceptable prediction is set at an extremely high level, namely, must be based on a governing regularity. Little imports both of these from the natural sciences into the social sciences.[12]

Little says that once the underlying causal structure or mechanism is known, then there is no longer a need for the phenomenal properties and that better predictions result when one replaces the phenomenal properties with governing properties. But, as was noted in Chapter 4 with respect to the hypothetical example of the Reagan administration's use of Wallace's 1982 study of arms races and war, until the latter are available, one should accept the former as enlightening and satisfying at the time it is offered (but not as the terminus of all enquiry). According to Little, in studying the social world one will not find governing regularities. So it makes sense to conclude that in the social sciences one will have to make do with the benefits, limited though they might be, of explanations based on phenomenal regularities. Thus these latter regularities may be regarded, at least to some degree, as enlightening and satisfying.

Little argues that some mistakes in understanding the social world come from the use of an analogy with the natural sciences. Consequently, many contemporary theorists have abandoned naturalism primarily because they find the conclusions of certain sorts of naturalists, namely empiricists and positivists, insupportable, not because of persuasive general objections specifically to the project of constructing a theoretical framework for the social sciences that contains some of the epistemic virtues of the natural sciences. One may not dismiss naturalism *a priori*; the project seems eminently worthwhile, as Chapter 1 showed. The crucial question is whether it can be executed. One must examine and evaluate the arguments for and against it on their merits.[13]

Little says that phenomenal regularities do support prediction, but they do so too weakly to be helpful. They are neither reliable nor capable of playing a role in the testing of social science theories. Phenomenal regularities are discoverable, but 'have little explanatory import'. Little is, at the very least, premature in this assertion. If these regularities are not reliable for prediction, then this needs to be established empirically. There are theoretical arguments against their reliability (e.g., Doran's below). But the claim that they do not have explanatory power for the reasons Little cites, such as extensive *ceteris paribus* conditions, is an account of *how* they might fail. *That* they fail still needs to be proven.

One last point is that Little's attack on social prediction is based on the claim, as noted above, that models and hypotheses make simplifying assumptions that are literally false and distort predictions.[14] But it is not clear why this should undermine the predictiveness of social science theories, since Nancy Cartwright argues that the same thing happens in physics. Yet, one might observe, the predictions in physics can be reliable enough to be of practical utility despite the distortions.

Little's attack on prediction is thus defective because, first, he mistakenly holds that no regularities stronger than Cartwright's weaker sort may hold in social science disciplines like IR. Second, contrary to what Little says, predictions based on phenomenal regularities should be admitted in the social sciences since, unlike in the natural sciences, there are no sounder foundations for prediction available to replace phenomenal regularities – especially in view of the practical requirements of prediction in policy-making discussed above. The claim that the regularities in the social sciences are insufficient is highly suspicious. It is much more plausible to claim that one will not find in the social sciences the same predictive ability that is to be found in the natural, or specifically, physical sciences. Neither Italian, French, Swedish nor Turkish troops will invade China in the next decade. Chinese leaders know this. China's military planning would probably be different if the Chinese government did not know this. This knowledge thus has practical value for the Chinese government.

Third, perhaps the binary distinction between governing and phenomenal regularities should be eschewed, as the relationship between a law-like statement and an underlying causal structure appears more complex in the social sciences, given that there seems to be more of a continuum than a neat twofold distinction that may be more appropriate in the natural sciences. Fourth, simplifications and distortions that Little observes to occur in social sciences also occur in physical theory, where their presence does not render them incapable of supporting prediction. Finally, Little's claim that phenomenal regularities have scant explanatory import cannot be settled *a priori* but must be shown empirically.

Little's argument against prediction shares with Bohman's the acceptance of causal mechanisms in the social sciences and the tenet that causal mechanisms are too weak (though each presents this in his own way) to support social science prediction. But one must ask of these two authors, are the minimum standards which they set for a proposition to qualify as an acceptable prediction too high? In both cases the answer is 'yes'. Little and Bohman, while generally rejecting positivism, seem unable to free themselves entirely from its grasp. Both appear to retain vestigial elements of the positivists' grandiose conception of a predictive statement, which has the character of a determinate or 'point' prediction. If one were to define more explicitly what is to count as a prediction, and if one brings the definition in line with what social scientists and policy-makers often take to be a prediction – in terms of the certainty or confidence level it must achieve – it may be possible to construct a justifiable notion of social 'prediction'. This is revisited below in the next section.

Non-linearities and the foundations of IR prediction

Opposition to prediction has come from several prominent power-transition theorists. Robert Gilpin has argued that while 'it is certainly possible to identify crises, disequilibrium and incompatible elements in a political system … it is most certainly not possible to predict the outcome' (1981: 47). He concludes that, while national leaders' beliefs about the future 'frequently become self-fulfilling prophecies', nevertheless, he insists, 'no one can predict the future' (1981: 232). Gilpin frequently disparages IR prediction, but does not offer a sustained argument for his scepticism. Another well-known power-transition theorist, Charles Doran, has provided the strongest anti-predictive argument in that tradition. Doran, in a recent work (1999), attacks predictiveness in IR and the social sciences by focusing on the non-linear nature of power transitions and supports his position with quite a distinct set of arguments, though the overall anti-predictive conclusion is one he shares with many authors in the mainstream of IR meta-theory and the philosophy of science. Like all contemporary scholars, Doran rejects explanation–prediction symmetry, arguing that scientific-style explanation in IR is possible but that prediction is not.[15] Doran does not come to meta-theory through the HT or indeterminism but through study of non-linearities in IR.

Doran's critique of prediction

Predictions about human behaviour, according to Doran, are based on laws that are usually probabilistic because they have such a multitude of causal underpinnings and exceptions. A complete prediction specifies what the event is, when it will happen, and explains how it will come about. Doran distinguishes several sorts of incomplete prediction: 'point', 'alpha' and 'beta' predictions. The first identifies only when (within a year) the event will occur, the second states what event will occur and when but lacks an explanation of how it will come about, while the third states what will happen but neither how nor when.

Doran notes that there are various sources of error in prediction, which include non-linearities, noise around the trend line, too few data points and long-term deviance from earlier trends that do not show up in the short term. Doran focuses his critique on the first, which he regards as the most serious and intractable.

States rise and fall in power relative to the other states in the system. Doran observes that the trajectory of the rise and fall tends to follow an s-shaped curve, which describes first the state rising slowly in relative power, then accelerating in relative growth, then slowing down in relative growth, and finally declining in relative power. The curve will contain several points of 'non-linearity', where, according to Doran, the 'tides of history' turn and 'everything changes'. At these 'critical points' the future no longer resembles the past.

There are two types of non-linearity. One type occurs when the curve goes from positive relative growth to negative relative growth or vice versa. The other type occurs at inflection points, where the torque of the series of tangents to the curve changes from anti-clockwise to clockwise or vice versa (see also Doran 1991: 97–100). Before the first inflection point, all straight-line extrapolations underestimate the future relative power of the state; after the first inflection point, all linear extrapolations will overestimate the future relative power of the state. The actual change in relative power at the inflection point is small. Predictions from straight-line extrapolations are made based on the evidence at the time of the extrapolation. The tangent to the curve at that moment describes the predicted growth (or decline) of the state. The predictions made at most points on the curve are fairly accurate, especially in the short term, but those made at inflection points are wildly inaccurate.

Doran illustrates his argument using the IR predictions of power-cycle theory but includes also the economic example of commodity speculation. He contends that non-linearities can never be predicted even though, in both cases, techniques and methods have been developed to refine statistical operations. All predictions will be made from patterns exhibited in the past but when the events develop in such a way (as they do when there are non-linearities) that what happens in the future is fundamentally different from what has happened in the past, the future events will be unexpected and unpredicted (Brown 1994). Predictions may be accurate as long as they are not made before the inflection point about an event after that point. A theory might be able to predict that a non-linearity will occur, but not when it will occur (Doran 1999: 21, 24). One of Doran's key critical arguments is that predictions are of little value in IR because they utterly fail at these critical points, just where they are most needed.

Dynamical systems modelling generates predictions in some domains. But Doran argues that one of the requisite conditions of application is that they model closed systems, a requirement that the world of IR does not satisfy. There will, moreover, always be a choice of different feedback systems that will be indistinguishable as proper models. The various models may be able to provide explanations of phenomena that have occurred. But since they will diverge on the question of future trends, it will be impossible to base predictions on them.[16]

Doran discusses 'power-cycle theory', which he has developed elsewhere (e.g., Doran 1991), and bases much of his argument on it. Doran says that power-cycle theory includes '1) the dynamic of state rise and decline itself' and '2) the implications of that dynamic for major war' (1991: 10). These tenets, along with the rubric 'power-cycle theory', would seem to suggest that Doran is drawing on the substantial literature of cycles and transitions. One might thus conclude that the 'predictive failure' variable is a standard part of that literature. This is not the case, which Doran points out (see 1989: 91–3 and 1991: 19–24, 95–100).

Doran suggests several ways that policy-makers can improve predictions: experimental refinements of linear forecasts, more frequent updating of the indicator data and predictions based on them, use of multiple indicators and continual improvements in information (1999: 36). Doran says that most of these 'are not likely to help' (1999: 34). They help, as noted, only when there are no non-linearities. But there is no way to tell in advance whether the predictions will cover a period of non-linearity. 'Better that [policy-makers] not forecast than that they naively expect success where there is none' (1999: 36). Doran concludes: '[i]n general … the capacity to divine the future is very limited' (1999: 37).

Doran finishes by offering a simile. A policy-maker guided by a theorist, who predicts by fitting theories to historical experience, is akin to a blindfolded driver taking directions on a winding, unmapped road from a navigator who only looks out of the rear window. The directions ('hold the wheel steady') may be good when the road is straight but they are unlikely to lead to a safe travel when there is a curve ahead.[17]

Analysis of Doran's account

Predictions in IR are at best imperfect and have inherent limitations. Still, they have value, which Doran underrates. The argument Doran offers against the value of IR prediction seeks to show that one will be correct most of the time when predicting that the near future will be like the recent past but these are the times when one does not really have much need for predictions (Doran 1999: 21). Rather, one needs reliable predictions most at those rare moments of non-linearity, when everything changes. Yet that is when they are not available (1999: 21). The non-linearities are especially important because historically those are the very moments when systems are unstable and war-prone. The core of Doran's critique can be found in the following two propositions.

(P1) *Correct predictions are possible when there are no non-linearities but predictions at these times have relatively little value, since they are not needed by states in order to plan policy.*

(P2) *If correct predictions were available when non-linearities occurred, states would be more secure and/or the system would be more stable.*

Both of these propositions seem, upon examination, to be false – or at least to stand in need of extensive qualification. Consider proposition P1 by imagining a state that operates each day in complete ignorance of the future, without any ability whatsoever to predict. That is, it has no expectations whatsoever about what will happen on the following day.[18] Is the state likely or unlikely to be confronted the next day with invasion, comprehensive trade embargo, revolution, etc.? Will the next day or next month be just like the last? The state needs to answer these questions in order to arrange its priorities in its allocation of resources. Yet, without rationally grounded expectations, the state would have no way to answer these questions. There is, then, considerable value to accurate prediction, if such a thing is possible.

One might respond to this argument, following proposition P1, that in times when the environment remains relatively unchanged, the state does not need prediction, since the state can assume that the environment will remain stable. On what is this assumption based? Why assume that the environment will be stable any more than that France will attack China tomorrow, or any other random claim about the future? If the answer has to do with the rationality of basing the belief about the future on a body of evidence, then one is engaging in prediction, given the definition on p. 8 (see also p.136 in this chapter).

If Doran believes that the state will generally continue to prepare for the future to be like the past, then the state continues to base policy on an expected future, i.e., on predictions. Those predictions are of great value, even if they are simple straight-line extrapolations of recent trends, because the state needs them to be able to prepare, however imperfectly, for the future. Straight-line extrapolations of the present are just as much predictions as any other sort of projection and over time and will permit more success in achieving foreign-policy goals than no prediction at all, or more success than taking seriously (and thus having to make preparations for responding to) random or supremely remote contingencies, such a possible near-term French invasion of China.

The point here is that even when things are moving smoothly, the state must make predictions about the near and medium-term future and base policy decisions on them. Policies will often be successful in part because the state has been able to predict more or less accurately (often using linear extrapolations). Hence, even though the state may experience predictive failure around non-linearities, the much-more-frequent predictive successes have significant value for the state, apparently much more so than Doran recognises, based on his endorsement of proposition P1.

According to the next part of Doran's critique, proposition P2, predictions lack significant value because at non-linear points when major war is most likely and states most need accurate predictions, accuracy is unattainable. Because Doran concludes (1999: 36) that is better not to predict at all than to do so naively as policy-makers do, his cost–benefit argument against prediction succeeds only if he can show that the results of prediction at non-linear points are worse than eschewing prediction at all times. This follows because if predictions are helpful when accurate, and if they are at least not harmful when inaccurate (at non-linear points), then the use of prediction would have overall positive value. Doran indeed argues quite explicitly that faulty predictions at non-linear points are dangerous. Because states cannot 'anticipate nonlinearities in the trajectory of their power cycle' they are 'unable to discount unpleasant or unfavorable events in advance' which in part provides an 'explanation for the onset of major war' (1999: 29).

How do bad predictions help explain war? By what mechanism do they cause instability? Doran (1999: 20) says that the analytical failures that states experience at non-linear points can occur in two different ways. Either (a) the state fails to predict or notice a genuine non-linearity in its power curve or (b) the state interprets a blip on its curve as a genuine and permanent change in its fortunes and the consequent overreaction adds to the 'degree of uncertainty' and 'belligerence' in the system. 'A government that always overreacts is a problem for world order' (Doran 1999: 20). So proposition P2 can be divided into two cases:

(P2a) *The state fails to notice a permanent shift*, or
(P2b) *A state belligerently overreacts to a temporary shift, which it mistakes for a permanent shift, which results from failed prediction.*

With respect to P2a, Doran holds that it is common for state A initially to miss the non-linearity (1999: 20). This oversight stems not from anything that state A in particular is doing wrong but from the inherent impossibility of predicting non-linearities. So state A's adversary, state B, will similarly miss A's downturn, since it is likewise impossible for others to predict or identify in advance the non-linearity – one's own or another's – in a state's trajectory. How, then, can this be a disadvantage for state A and how can it lead to otherwise-unlikely major war?[19]

Presumably either A initiates war or some other state, B, does. If A initiates war, then it would seem that A was not caught off guard by B (i.e., did not suffer as a result of its failure to predict); so predictive failure seems an unlikely causal factor. If B initiates war, then the question to ask is, 'How was state B able to see the non-linearity coming while state A did not?' There is no basis for an answer on Doran's account, since A and B share an inability to predict when a non-linearity will occur, as just noted. State B has no more chance to make preparations to take advantage of it than A has to prepare itself for B's actions. The onset and outcome of the war seem to be a result of the shift in the relative power of the states, not of the ability to predict and prepare for war in advance.

The most well-known power-transition theories, like that of Organski and Kugler (e.g., 1980), offer a rational actor, utility-maximising explanation for the instability at transitional periods of a system. For example, the rising state has a newly-found ability to challenge the hegemon (or dominant states), which has structured institutions to its (or their) benefit, or the hegemon has the incentive to pre-empt and undercut forcibly the challenge posed by the rising state. Doran's power-cycle theory adds another element to the causal account of major war, the state's inability to predict non-linearities on its power curve (see Doran 1999: 29, cited above). He says, 'failure to predict is a determinative component of power cycle theory' (1999: 29). Later Doran says, 'According to power cycle theory, it is the absence of predictability … that accounts for … the largest, most severe wars of history' (1999: 29–30). But the problem of unpredictability (or the suddenness of becoming aware of changing power relationships) that brings about war does not appear among the basic terms of the most well-known power-cycle theories, such as those of Gilpin (1981, 1989), Kugler and Lemke (1996), Kugler and Organski (1989), Organski (1958), Organski and Kugler (1980), Modelski (1987) and Modelski and Thompson (1989). While this study does not endorse any current power-cycle theory, it recognises that the causal mechanisms usually cited have *prima facie* plausibility and do not engender any obvious inconsistencies.

Power-cycle and power-transition theories are plausible and no attempt is made here to discredit them. But their *prima facie* appeal should not lead one to regard 'predictive failure' as a valuable causal variable. Are the causal mechanisms Doran endorses, including 'predictive failure' leading to war, as appealing as those of power-cycle and power-transition theories?

Suppose, contrary to Doran's claim, that a state could accurately predict an inflection point on its power curve approaching in, say, three years. What would the state do differently? It cannot change the fact that the inflection point is coming, according to Doran, since he says explicitly that 'there is relatively little any government can do about changing latent power in the short term' (Doran 1999: 36). Doran (1999: 29) also adds that, according to his power-cycle theory, some states act belligerently when they perceive an inflection point. So if states cannot prevent a non-linearity from occurring, and if some states take advantage or act belligerently when they perceive non-linearities, it is difficult to see, at least on Doran's power-cycle theory, how the ability to predict these fundamental shifts in the balance of power could prevent war.[20]

Doran says that states cannot circumvent non-linearities in the short term but does not say whether more can be done in the long term. Let us assume that indeed more can be done to change the future, if a state has several decades to prepare. The failure of prediction (understood as it is by Doran as 'point-prediction') might not constitute a serious problem because Doran says that states cannot predict the non-linearities with accuracy within a year or two. States plan far into the future. For example, the tens of billions of dollars authorised for an aircraft carrier is expected to continue contributing to the state's security thirty years after authorising the investment. So when a state is planning several decades into the future, it is much less important to be able to determine in precisely which year a future event will occur, since long-term planning does not require such point predictions. P2a above, which deals with a state's failure to notice a permanent shift, does not seem to be a plausible claim, given that states cannot do much to alter upcoming non-linearities, even if they could predict them in the short term. Moreover, it does not follow from Doran's analysis that long-term predictions, with much wider time-frames, are impossible, given the current state of IR theory.

What then about P2b, which deals with the effects of mistaking a short-term (downward) shift for a long-term shift? Would major war be less likely if erroneous prediction were eliminated? According to Doran, some states act belligerently when they mistakenly perceive a non-linearity. Doran also acknowledges, as one would expect, that when states correctly perceive a non-linearity they 'may begin to act belligerently … precipitating war' (1999: 29). Because power-cycle theory says that those downturns will come about sooner or later, it follows that eliminating incorrect prediction would not eliminate major war. Would the system be more stable if non-linearities were correctly predicted? It seems not, since Doran suggests in P2 that predictions are most dangerous not to have when they are most needed, i.e., when real discontinuities occur. That suggestion must now be understood within the context of a theory that includes the tenet that even if the discontinuities were correctly predicted, war would usually still result.

To put it another way, the unpredictability of non-linearities would seem to affect only the timing of war, not its occurrence. First, as noted, Doran says that states can do little to change relative power trends. Second, Doran (1991: 94–5) emphasises also that states quickly see critical points when they arrive. Third, if non-linearities were predictable, they would be predictable by many parties, since A's decline in relative power is often not a result of actions by or trends in A, to which A would have privileged access. Rather, as often as not, A declines relative to others because rival states are increasing their absolute capabilities faster than A (see Bueno de Mesquita 2002). So the conclusion that seems to follow is that if non-linearities were predictable, then an event (war) that occurs at any non-linear moment would simply occur at an earlier point when the prospect of the non-linearity came into focus in the analysis-cum-prediction of the major states. Even if it were possible to predict non-linearities, on Doran's power analysis that ability to predict would not help avert war. Consequently, the inability to predict cannot be regarded as a causal factor in the onset of war.

One further consideration is important in the evaluation of the second part of Doran's anti-predictive argument. Doran says that predictions are most needed at non-linearity points and that they are unavailable then. But are they? Perhaps some predictions will fail then, perhaps many more than at other points in the state's history, but many predictions will remain reliable. Doran says at the non-linear points the tide of history turns and 'everything changes'. Does everything change? Granted, many more things than usual change and the slogan 'everything changes' expresses an important idea. But in the process of the cost–benefit analysis of prediction, it is important to know how literally true this is. The answer is that, much changes, but much does not and many predictions remain reliable.

Consider again two members of the UN Security Council, France and China. Suppose China is entering a non-linearity. It would be important to know whether France is likely to invade the Chinese mainland in the next planning period, say the next five years. Both states are major world powers and China would have to be concerned if there was a chance that a large state with a military force of 350,000 soldiers and nuclear weapons were likely to attack. But the prediction that France will remain at peace with China in the next five years is an extremely high-probability prediction.

Imagine that this year China is entering an inflection point. It remains true that France will not invade China in the next five years (even if France were also entering an inflection point). If France were likely to invade China, Beijing would want to re-evaluate its diplomatic strategy towards the West and towards its neighbours, its military procurement and force structure policies, and the distribution of its intelligence-gathering assets. Certainly East Asian states have fought wars against Western states; so in general such attacks are conceivable. Nevertheless, China can predict continued peaceful relations with France. That China knows that it will remain at peace with France for the next five years, even if both are entering inflection points in their national power curves, is non-trivial knowledge for China. Many different theories that seek to predict would predict France's non-aggression toward China, and they would be reliable predictions whether or not inflection points are imminent.

To return to Doran's example of a blindfolded driver being directed by a navigator looking out the rear window, the driver needs to know when it is likely to be beneficial to keep the steering wheel straight. Crashes at curves might be likely. But crashes on the decades-long stretches of straight road can be avoided if the driver is correctly instructed to keep the wheel straight. It is just as helpful to avoid crashes when the road ahead is straight as it is when the road ahead is curved. And it is better to avoid crashes on straight stretches of the road and run the risks on curves, than to risk crashes both on straight and on curved sections of the road because of a complete lack of information. Since neither proposition P1 nor P2 seems sustainable, neither is Doran's conclusion downgrading the value of prediction in IR.

Prediction and probability

The development of a successful account of 'prediction' that avoids the criticisms of the authors discussed above will require important use of the notion of 'probability'. By incorporating the notion of 'probability' it may be possible to modify Doran's position to provide a basis for prediction. Doran observes that regularities in the social sciences are probabilistic. He also mentions the notion of 'confidence' (1999: 13), which he equates with 'reliability'. One may have greater or lesser evidence for a particular probability statement. Progress may be made by considering the distinction between the confidence value of a statement and its probability value. It is possible to calculate the probability of a future event, and separately estimate the degree of confidence appropriate to that probability assignment by taking into account where the state is on the power-cycle curve.

Doran argues that states' power curves follow a regular pattern, the s-shaped curve. Leaders, if they were to adopt Doran's analysis, could adjust confidence levels as they move along the curve. For example, the longer a state's relative power accelerates, the less confidence one should attach to predictions. This would then rationally require more consideration of and planning for alternative futures. States do this all the time, since predictions are not generally deterministic, wherefore planning is based on alternative (non-equiprobable) futures. The lessons of power-cycle theory (and the same could be applied to theories that accept other regular patterns or cycles, e.g., long-wave theory) would be to adjust the confidence of predictions based on historical patterns. Thus by distinguishing probability from confidence, it may be possible to move toward a partial resolution of the policy-maker's dilemma with which Doran concludes.

Prediction-scepticism results in part from overlooking the importance of 'probability' in hypothesis-formation. Little (1991) talks about probability statements, but not in the context of prediction. Bohman has even less to say about probabilistic prediction. This lacuna weakens their positions. Part of what leads them to their anti-predictive conclusions is the high, deterministic-style standard they set for prediction. While leaders may not see the world as fully indeterminate in the way that Habermas, Bohman and the HT do, they see it as a world of habits, tendencies and probabilities. Leaders regard prediction as a matter of probability. (One of the most well known is President Kennedy's probability assignment of war with the USSR during the missile crisis, see Sorenson 1965: 705). Once this feature of the thinking of decision-makers is recognised, it is possible to combat some of the confusion generated by post-positivist meta-theory.

Rational choice and Bayesian decision theory require not only a utility assignment for each outcome but also a probability assignment. If one believes that phenomenal laws justify probabilistic predictions, then there is a basis for policy decisions. While it is clear that states and leaders typically view future outcomes as a matter of probability, it remains to be settled whether 'probability' should be treated as probability1 (an empirical notion of 'relative frequency') or probability2 (a logical notion of 'degree of confirmation') as Carnap (1950) distinguishes them, or whether one of the other interpretations should be preferred (e.g., Popper's propensity theory, or DiFinetti's (1964) subjective assignments). But historical examples make quite clear that leaders typically regard future outcomes not as determinate but as matters of probability, which highlights Little's and Bohman's over-estimation of what it takes for a proposition to qualify as a 'prediction'.

States and leaders must make decisions. Leaders who seek to bring about outcome T will consider policies, Q, R and S (where Q, R and S may be complexes of specific actions – e.g., unilateral arms embargo against state A and alliance with state B – and T may be a complex mix of specific goals – e.g., mixing minimum levels of prosperity and security). Where '$\rightarrow$' is 'brings about', they will evaluate $p(Q\rightarrow T)$, $p(R\rightarrow T)$ and $p(S\rightarrow T)$.[21] Whichever of the three has the highest value will be the chosen policy. Even if the values of the three are relatively low, say, 0.3, 0.2 and 0.1, respectively, with similar confidence levels and expected costs, it is still rational for the leader to choose Q because it is more likely than any of the others to lead to the desired outcome. A rational leader would choose the policy option that maximises the chance of the desired outcome, though there would be more resources committed to planning for other contingencies (hedging of bets) than if the highest choice probability were 0.8 or 0.9. A weak rational basis for policy choices is preferable to random action, which has no rational basis at all. Hence, given the best available theory, a weak theoretical link between action and desired outcome is better than no link at all.

Doran's prediction-scepticism argument provides an important insight into prediction in IR. While it is argued here that Doran's general scepticism does not follow, two significant conclusions may be drawn: decision-makers must attend to confidence levels of predictions and predictions will fail at non-linear points on the power curve. The first point is that, because predictions will fail at non-linear points, policy-makers must be conscious of their possible failure at those points or in those intervals. If one accepts Doran's argument about the typical power curve and the existence of non-linearities, then it follows that there are limitations on prediction, since the points of non-linearity cannot be identified in advance. However, using the framework Doran outlines and tracking the power trajectory of the state should at times suggest entry into a period of greater suspicion that non-linearities may occur. At such times there should be heightened caution in the sense that confidence in probabilities assigned to future consequences is reduced, and there should be a commensurate increase in contingency planning for alternative outcomes and in the hedging of one's bets.

Even if one rejects the idea that there can be a rational basis, built on Doran's own power-cycle theory, for being more suspicious of impending non-linearities, there are reasons to reject Doran's prediction-scepticism. One may grant that predictions will fail at non-linear points. But almost all of a state's history is at points other than those Doran emphasises. While it may be true that those points are the most important in the state's history, it was argued above that the welfare of the state is much better served by making correct (or usually correct) predictions for years at a time and then suffering the consequences of failed predictions (especially if greater caution attends those predictions during suspicious periods) than would be the case if prediction were abandoned altogether. If Doran were to respond that at all times there is an equal chance of non-linearities, the same argument against his scepticism could be offered in a modified form, by admitting that confidence levels should be suppressed all the time. But the fact that well-constructed predictions are still reliable 99 per cent of the time is a powerful reason to continue to base policies on prediction, even if all the time it will be necessary to hedge bets and devote more resources to the non-predicted contingencies. And, as noted, rival states will be just as likely to fail to predict a given state's non-linearity.

Bernstein, Lebow, Stein and Weber's critique of prediction

The discussions of non-linearities, conventionalism and probabilities are helpful for the task of evaluating the criticism of prediction presented by Bernstein, Lebow, Stein and Weber (2000), who argue that there is room for naturalism. As Chapter 2 shows, for much of the twentieth century the dominant position in the philosophy of social sciences was naturalism, namely, the view that the natural sciences constitute a model of knowledge that the social sciences should adopt. The particular form it generally took was an empiricist-positivist position – what Salmon (1989) calls 'the received view', according to which classical mechanics is the model of scientific knowledge that all branches of knowledge, including the social sciences, should strive to emulate. Newtonian physics is useful, in part, because it allows prediction of any future event just as easily as it allows explanation of any past or present event. And as Chapter 1 argues, makers of both foreign and domestic policy must be able to predict.

Over the past thirty years the empiricist-positivist 'received' view has come under sustained criticism and has been largely discarded, especially as reflectivist and hermeneutic principles came to be increasingly adopted by philosophers of social science. The 'received view' included the d-n tenet of 'explanation–prediction symmetry'. Once that tenet was disproved, prediction was rejected in favour of explanation.

In recent years those who have endorsed naturalism have generally defended it by arguing that the empiricist-positivist paradigm of classical mechanics is not as good a natural science model as some alternatives. Bhaskar and his followers (discussed in Chapters 3 and 4 – and in IR see especially Patomäki 2002) reject the empiricist-positivist element in favour of a non-empiricist 'critical realism', in which the parallel between the natural sciences and social sciences obtains. Bernstein et al. defend naturalism by arguing that the natural science model for the social sciences is not classical physics but rather evolutionary biology.[22]

Bernstein et al. categorically reject the predictiveness of international theory. They say '[i]nternational relations scholars cannot predict the future…' (Bernstein et al. 2000: 52) and '[t]he future is not predictable' (Bernstein et al. 2000: 70). Bernstein et al. justify their 'pessimism, both conceptually and empirically, and argue that the quest for predictive theory rests on a mistaken analogy between physical and social phenomena' (Bernstein et al. 2000: 44). In their view, '[e]volutionary biologists do not aim at prediction but instead have focused their efforts on developing theories that explain the process and history of evolution' (Bernstein et al. 2000: 70) and that 'the study of evolution … [through the] scientific approach should be of particular interest to political scientists because it eschews prediction in favor of explanation' (Bernstein et al. 2000: 49). In sum, they say their view is that, in comparison with physics, '[e]volutionary biology is a more productive analogy for social science'.

Although they disparage prediction in IR, they do so in a milder way than authors like Doran or thorough-going reflectivists like Patomäki. Part of the reason for their reluctance to embrace more radical scepticism about predictiveness may be that they begin with a discussion of the history and original purpose of the academic discipline of IR, which was tied to policy needs – of preventing war and making the world a more peaceful and prosperous place. Thus, as Chapter 1 showed, serious consideration of the history and original purpose of IR makes it much harder to ignore policy-makers' needs and the requirement of prediction (see also Chernoff 2002).

Most of the authors who attack 'predictiveness' of social science theory do not offer a clear definition of the term. Doran is one of the few authors who does define the term. The present discussion makes use of the definition of 'prediction' on p. 8. The justification of prediction will require the use of theory. The policy-maker must judge what results are likely to follow from the choice of a particular policy, which cannot be done without having a general set of principles that offer such links, i.e., without adoption of cause-and-effect linkages, which are produced by theory. Policy-makers must make decisions and decisions require beliefs about what developments will result from each of a set of policy options. Such beliefs about the future are, according to the definition on p. 8, predictions. Bernstein et al. attack the Newtonian analogy, argue that IR theory is more like evolutionary biology than classical mechanics, and offer 'scenario analysis' as an alternative to replace theory-based prediction in IR, which are treated in the next three subsections, respectively.

The Newtonian analogy

Bernstein et al. attack the Newtonian analogy by arguing that IR encounters un-Newtonian-like problems of:

1 definition, measurement, coding
2 uncertain relationships between underlying and immediate causes
3 learning
4 the single-case problem and
5 the non-linearities of open systems.

The following sections assess arguments 1–4 of Bernstein et al. Their fifth criticism has already been examined above in the course of the appraisal of Doran's much more detailed statement of it. After examining the critique of 'theory-based prediction', this section then examines the alternative that Bernstein et al. offer to traditional IR theory-based prediction, namely, scenario-construction. Much of what Bernstein et al. say about scenario analysis is fundamentally sound and quite illuminating. What is questioned here is that their method is not an alternative to the approach of standard IR theory but rather is an instance of the application of IR theory and that it relies on the probabilistic conclusions IR theory can produce.

Bernstein et al. unduly diminish the value of prediction. They argue that IR is very different from Newtonian mechanics and is much more like evolutionary biology. The former is predictive, while the latter is not. Hence IR should be viewed as non-predictive. On their view IR is 'scientific', since biology is certainly a science, nevertheless, IR is not predictive because of the profound differences between classical mechanics and IR theory. This section will show how the sharp divergence between Newtonian mechanics and IR is overstated, primarily because Bernstein et al. offer an oversimplified account of classical physics.

Definition and meanings of theoretical terms

How one understands the functioning of the terms of a theory has a significant effect on how one conceives of the character of theories. Chapter 3 attempted to lay out a distinction between two types of terms in empirical sciences, observation and theoretical terms. Bernstein et al. argue that there is no consensus on the meanings of key terms in IR, which is a result of 'the arbitrary nature of the concepts themselves' (2000: 46). In contrast, terms in physical science 'are embedded in theories with deductive implications that have been verified through empirical research. Propositions containing these terms are legitimate assertions about reality because their truth-value can be assessed' (2000: 46). This characterisation of the sharp difference in the functioning of theoretical terms between the natural sciences and social sciences greatly overstates the differences. For example, in the philosophy of natural science there are substantial differences over even the most fundamental questions, such as whether theoretical terms refer to real objects. As noted in Chapter 2, scientific realists like Boyd (1973), the early Putnam (1975), and Leplin (1997), hold that the terms of scientific theories denote objects in the real world, while non-scientific realists like Van Fraassen (1980, 1989) and Laudan (1981) hold that they need not be construed as referring to such.

In IR, one of the core disagreements between neorealists and neoliberal institutionalists is over the role and influence of international institutions in international politics. In the past decade there has been disagreement over what can be concluded about IR theory from the observable behaviour of NATO, which seems to have a good deal of influence in European and world affairs, and which could be viewed either as a realist-style alliance or as an international institution. Is the disagreement over NATO a result of irrevocably incompatible definitions alone – due to the 'arbitrary nature of the concepts'?

Consider the neorealist definition of 'institution' by Mearsheimer (1994–5) and the neoliberal institutionalists' definition by Young (1982). Mearsheimer says that an institution is, 'a set of rules that stipulate the ways in which states should cooperate and compete with each other' (1994–5: 8). Young, in the course of defining the term 'regime', says that regimes as well as 'all social institutions … are recognized patterns of behavior or practice around which expectations converge' (1982: 277).[23] As one can see, NATO in the twenty-first century appears to qualify as an 'institution' under both definitions. So the disagreement is over something more substantive and less arbitrary than simply two groups of theorists who define individual terms to refer to disjoint sets of objects. So while neorealists and neoliberals disagree over what NATO shows about the roles of institutions in world politics, they do not disagree that NATO (or the UN or the EU, for that matter) is an institution, given the above definitions. Mearsheimer's and Young's definitions are not identical, but their differences do not lead to different conclusions over whether NATO or any of the institutions that are at the centre of the debate qualify as institutions.[24] The nature of the concepts is not so arbitrary that the two sides classify NATO in different categories. The disagreement is not a result of 'arbitrariness of the concepts'.

This is not to say that theoretical disputes never turn out to be a result of incompatible definitions; often they do. But these are not insuperable problems and are not a result of 'arbitrariness' of the concepts that infects social science but not classical mechanics. The function of terms is more similar in the two areas than Bernstein et al. acknowledge (as noted on p. 79). In the seventeenth-century debate over the nature of 'motion' and the principle of the conservation of motion, part of the disagreement stemmed from the central importance physicists placed upon deriving the simplest possible theory. Physicists agreed that it was simpler to conceive of 'motion' as conforming to the principle of conservation than otherwise. Newton's definitions of 'motion' was chosen over Descartes' in large part because it fulfils the 'conservation' desideratum. The meta-theoretical principle of 'simplicity' was used to help define a key term.

With regard to their critique of conventionalism, Bernstein et al. contend (2000: 46) that 'propositions containing these terms are legitimate assertions about reality because their truth-value can be assessed' in a more rigorous way than the propositions of the social sciences. How have physicists determined the truth values of propositions employing these terms? In the case of the observational term 'motion', as noted, they have done so by appeal to the extra-theoretical principle: accept the simpler among competing theories, other things being equal.

One should also note that the version of conventionalism defended here holds that in all physical theory there is a degree of conventional choice on the part of scientists. The core claim of conventionalists, that there are extra-theoretical, conventional choices that must be made in selecting a physical theory, is inescapable and is today recognised by all physicists. Although they are extra-theoretical, Duhem and others have argued that these principles, like that of 'simplicity', are non-arbitrary and rationally grounded. In the debate over the true nature of physical space, the Newtonian view posited Euclidean space, while relativistic physics saw space as curvilinear. Did the experimental observations that led to the acceptance of relativity theory prove that space is non-Euclidean, as the remarks of Bernstein et al. would seem to imply? Indeed not. Physicists have come to acknowledge that even the most widely accepted physical theory cannot guarantee (whatever the present state of empirical data collection) that no other theory will account for those data as well as the preferred theory. One of Duhem's most well-known and widely accepted arguments shows that there are always other theories that fit with all available observations as well as does one's preferred theory – the so-called principle of underdetermination of theory by data. Physical theory includes an axiomatic system of geometry – but physical theory and observation cannot guarantee that the system of geometry accepted is in any absolute sense the true geometry of the world. There will be an inescapably *conventional* element to the system of geometry physicists choose.

One wishes to answer the question, 'Is physical space Euclidean?' Perhaps one can measure space. But the answer will depend upon how the measurements are carried out. For example, a measuring rod may be used. But one may wonder, thanks largely to the parable offered by Poincaré (1905) to illustrate this point, whether the measuring rod maintains a constant length or whether there are forces acting upon it when moved that affect its length. Of course, the problem is not solved by introducing a second rod to measure the first, since the same forces will presumably act upon the second rod. So there is no way to prove that the measuring rod is rigid. The physical theory that is accepted will depend upon extra-theoretical principles, such as the so-called measure-stipulation e.g., that the measuring rod is of constant length. As noted above, physicists accept the principle that the physical world is simple. Other things being equal, they prefer simpler theories to complex theories. That principle helps them select relativistic physics over classical physics because even though non-Euclidean geometry is not as simple as Euclidean (as Poincaré argued before special relativity), the whole system of physical-theory-plus-mathematics is simpler in the relativistic framework. But the terms and propositions of physical theory are not tested based on some direct, self-evident methodology. There are contestable choices that must be made in the process, which is much more like the social sciences than Bernstein et al. acknowledge. Again, Bernstein et al. overstate the difference between physical theory and IR theory.

With regard to laws and idealisations, Bernstein et al. say that terms in physical science 'are embedded in theories with deductive implications that have been verified through empirical research. Propositions containing these terms are legitimate assertions about reality because their truth-value can be assessed' (2000: 46). They continue, '[s]ocial science theories are for the most part built on "idealizations"...' So in their view there is divergence between the natural and the social sciences in that only in the former do the best theories have genuinely true laws. Social science theories have laws that are at best idealisations or approximations because the social universe is so complex and has so many variables that come into play across a range of cases.

Philosophers have, however, long argued that the best and most acceptable laws in the natural sciences are literally false (pp. 96–8). Descartes endorses idealisations that do not, in a literal sense, express truths, arguing that false suppositions are sometimes useful for science (Clatterbaugh 1999: 58). The persuasive arguments of Cartwright (1983: e.g., 81) conclude that explanation should not tie laws directly to reality. The fundamental laws of the theory should be understood to apply only to the objects in the model, not to the real-world phenomena. The objects in the model have only the form or appearance of the real objects without having all of the proper qualities. For Cartwright in (non-causal) theoretical explanation, the explanatory law may very well do its work and succeed in meeting proper criteria for acceptability among laws, despite the fact that the law is not true (see pp.77–8).

In the social sciences the explanations of state behaviour in rational-choice theory, which posits the existence of a rationally behaving state, or in economics, which posits the existence of 'the rational economic person', are clearly idealisations. However, these are parallel to – and not completely disconnected from – the postulation of laws governing frictionless machines or ideal gases in the physical sciences.

With regard to underlying and immediate causes, Bernstein et al. advance a number of points in the course of offering the critique of 'predictiveness' of IR theory and the scenario-based alternative that Bernstein et al. present. The main point to be made with respect to the use made by Bernstein et al. of the 'uncertain relationships between underlying and immediate causes' is that, in their view, such a connection 'makes point prediction extraordinarily difficult' (2000: 47). As will become clear, this chapter argues that IR theorists offer non-point predictions and that such predictions are often of genuine policy value. A statement like 'a rapid withdrawal of US military forces from Iraq after the defeat of Saddam Hussein is likely to lead to increased violence and long-term instability', would not qualify as a point prediction. Yet it is the sort of prediction that policy-makers typically rely on and it has value for policy-makers.

Bernstein et al. thus commit a straw-person fallacy, as do other critics of 'predictiveness' discussed above. They raise the bar of 'prediction' much too high. The sort of prediction that policy-makers need is often much less than 'point prediction'. One may grant Bernstein et al. all of their points seeking to undermine point prediction and still conclude that they have not provided good reason to reject many sorts of statements that satisfy the requirements of prediction, as laid out in the definition on p. 8.

With regard to learning, IR differs from predictive natural sciences in that, according to Bernstein et al., '[m]olecules do not learn from experience. People do, or think they do. … We know that expectations and behavior are influenced by experience, one's own and others' (2000: 47). They argue that US policies vis-à-vis the USSR were a response to the failure of appeasement in the 1930s, and those policies were a response to British leaders' belief that more aggressive policies failed to keep peace in 1914. They cite concepts like 'chain reactions' and 'contagion effects' to describe these phenomena and hazard analysis for their measurement. But they charge that these do not succeed in explaining 'how and why these patterns emerge and persist' (Bernstein et al. 2000: 47). They note also that theories that attempt to predict the future predict incorrectly in part because groups (often states) react in such a way as to prevent their predictions from obtaining. 'Human prophecies … are often self-negating' (Bernstein et al. 2000: 52). Another approach that has considerable explanatory success is cybernetic theory, which lays out mechanisms and may even be viewed as predictive. Cybernetic theory was developed by Wiener (1949) and applied to IR especially by Deutsch et. al (1957), and specifically to foreign-policy decision-making by Steinbruner (1974), Chernoff (1995) and others. It not only accords with the data but offers a causal mechanism to account for the patterns.

Bernstein et al. argue that actors can change 'the rules of the game' and consequently 'general theories of process in international relations will have restricted validity' (Bernstein et al. 2000: 52). Generalisations will apply only to 'discrete portions' of history. They add that 'scholars need to specify carefully the temporal and geographic domains to which their theories are applicable. We suspect those domains are often narrower and more constrained than is generally accepted' (Bernstein et al. 2000: 52). This is, however, just what one of the most systematic and generalising of all IR scholars, i.e., Waltz, demands. He maintains that a body of propositions does not even qualify as a theory unless it so specifies. Once these restrictions on domain are specified, theoretical generalisations might have value. For example, Waltz says that whether we are interested in natural sciences or social sciences, '[n]o matter what the subject, we have to bound the domain of our concern, to organize it, to simplify the materials we deal with, to concentrate on the central tendencies, and to single out the strongest propelling forces' (1979: 68). He later adds, '[t]o be a success … a theory has to show how international politics can be conceived of as a domain distinct from economic, social, and other international domains that one may conceive of' (1979: 79). So it is not accurate to suggest that IR theorists, at least careful ones, do not bound or constrain the scope of their studies and generalisations.

While bounded theories may offer predictions, the fallible nature of all empirical knowledge and the probabilistic nature of IR laws does restrict what can be predicted. The longer the chain of reasoning and the greater the number of probabilistic propositions that are conjoined, the less one may rely on the predicted event. But this is an argument for limitations on certain types of predictions, not an argument against the predictiveness of IR or social science theory.

Bernstein et al. note also that there is the problem of the 'single case'. They correctly observe that policy-makers often worry about a single instance of an event type (e.g., possible war with our neighbours on our western frontier in the next year) rather than with general propositions (e.g., the problem of war in general, or great-power war, etc.). First, with regard to the relevance of theory to policy, it is important to note that policy-makers *do* sometimes care about long-run patterns or generalisations in some of their decision-situations. Long-run generalisations are very frequently important for policy-makers.

For example, in the 1990s, numerous new states emerged in Europe and central Asia due to the collapse of the Soviet Union and the end of its domination of central eastern Europe. Policy-makers had to understand as correctly as possible the truth of generalisations in DP theory in order to decide how much in the way of financial and diplomatic resources should be committed to help promote democratic polities in the region. Policy-makers in 1994 were not primarily worried about a specific war, say, between Hungary and Slovakia, but rather the long-run chances of conflict arising in a Europe with many non-democratic states as compared to a Europe with very few such states. Policy-makers could consider two alternative scenarios of the future. One is a Europe with some democratic states and many non-democratic ones, possibly including a non-democratic Russia. The other one envisions a Europe with virtually all democratic states. The generalisations about the effects of democracy on behaviour would allow policy-makers to draw general conclusions about peace and international stability from each scenario.

Second, whether there is truly a 'problem' is a function of the sort of understanding one has of probability statements. Different theorists have advocated different ways to interpret probability statements. As noted above (p.99), Carnap (1950) argues that no single interpretation is adequate for all probability statements and that some must be interpreted as expressing relative-frequency propositions and other as partial-entailment (hence purely logical) propositions. There are also subjective degree-of-belief and real propensity possibilities that have been advanced. Under some of the possible interpretations there is no particular difficulty with single cases. The single-case problem is most troubling if one takes all statements of probability to represent the long-term limit of relative frequencies. But it is not a problem if probability statements are interpreted along the lines of subjectivism, partial entailment or real propensities. Chapter 4 argued that at least a subjective or degree-of-belief interpretation and relative frequency interpretation are needed and that observations of long-run frequencies constitute *evidence for* probabilistic predictions. But that the *meaning of* those statements is given by a subjective or partial entailment interpretation. On this view the problem of applying them to single cases is minimal.

Scenarios, probabilities and prediction in international relations

The five principal arguments discussed above that Doran and Bernstein et al. offer against prediction in IR are flawed. Bernstein et al. proffer what they regard as an alternative, since they correctly observe that the policy-maker must be able to plan for the future, even though they deny that prediction is possible. They say, 'international relations scholars cannot predict the future, but neither can we ignore it. People need to make decisions in the face of uncertainty about the future, and consequently they need appropriate concepts and foci for information to maximize the quality of those decisions' (2000: 52).[25] Deductive-nomological modelling, according to Bernstein et al., is of 'very limited utility'. So in order for IR scholars to have any effect on policy, other methods of analysis must be devised.

The alternative method Bernstein et al. advocate, scenario analysis, has great merit and is close to types of analysis used by policy-makers in many countries. Bernstein et al. lay out the individual steps of scenario-writing to illustrate the problems of decision-making under uncertainty and some of the specific possible cause-and-effect sequences. Their recommendation that the scenario method can be used in conjunction with other tools is well taken (though they do not include predictive IR theory). However, the claims that Bernstein et al. make on its behalf are not entirely justified and, in particular, the contrast between it and theory-based prediction is not as stark as they portray. It will be argued here both that the criticisms of 'predictiveness' fail and that the solution Bernstein et al. offer is compatible with 'predictiveness'.

Throughout their discussion of scenarios Bernstein et al. (2000: 27) make four key claims, all of which are dubitable. The three points may be stated as follows: (i) d-n and theory-based prediction is point prediction, while scenario analysis aids policy-making without predicting the future in an unwarranted way; (ii) theory-based prediction entails claims of 'knowing the future', while scenario analysis allows for uncertainty and thus works with the unfolding of new information as the policy-interactions proceed; and (iii) scenario analysis allows the use of the most plausible assumptions, which can be revised as time passes and new information arises, while it avoids unwarranted theoretical prediction. They contrast IR theory and prediction, which show the authors' acceptance of a fourth claim, (iv) d-n-based prediction and the method of scenarios are fundamentally different from one another and only the latter can justifiably be employed by policy-makers.

With respect to (i) theoretical prediction does not rely on point prediction in some way fundamentally different from the method of scenarios. Not all IR theories claim to generate point predictions, since many permit prediction that does not meet that standard. Moreover, predictions may have clear policy value even though they are not point predictions. First, Bernstein et al. offer no explicit definition of 'point prediction', though presumably what they have in mind is something like the notion defined by Doran (p.143 above). While they use the terms 'prediction' and 'point prediction' in what, on the surface, seem like different ways, when they come to the specifics of what IR theory does, they say that it must, if it is predictive at all, seek point prediction and they do not leave room for some less precise form of prediction. Is all IR theory prediction point prediction? Theorists like Waltz certainly do not offer point prediction – a feature of his brand of neorealism some have regarded as a failing. Waltz explains trends and enduring regularities rather than precisely defined, individual events. Likewise, neoliberal institutionalists try to explain general trends and patterns without offering point predictions (Keohane 1984, Ruggie 1996).

Second, the discussion of the single-case problem above shows that general propositions, if justified, may offer an important ingredient to the policy-maker's analysis, even without point predictions. If a policy-maker knows that democracies are very unlikely to go to war with other democracies, then, even if there is no concern about a war this year between two specified Central European states, the policy-maker knows that it is worth putting resources into promoting more democracies in the region to reduce the incidence of war in the future. Various general statements about democracies, non-democracies and war are enough to justify the general (non-point) prediction that a Europe with no non-democracies will be more peaceful than a Europe with an admixture of democracies and non-democracies. And the latter is enough to justify certain policy decisions.

Bernstein et al. implicitly demand that if IR theory is predictive, it must generate point predictions and, in so doing, they seem to be raising the bar unjustifiably high by defining 'prediction' in an unnecessarily restricted way in order to ensure that standard IR theories do not succeed in clearing it. Why not compare scenarios to the predictions of IR theories that are *not* point predictions, like the statement 'Multipolar systems will exhibit more great power war than bipolar systems', or 'Democracies rarely go to war against one another'?

It is also interesting to observe that when Bernstein et al. are being most precise and most critical of what IR cannot do – in contrast to their praise of what Newtonian physics can do – the authors (and other critics of prediction in IR) often use the term 'point prediction' rather than just 'prediction'. Is there a difference between 'prediction' and 'point prediction'? As noted, Bernstein et al. offer neither a definition nor an example of 'point prediction'. Presumably, an example would be something like (A) 'Bilateral war between China and France will occur in 2013', or (B) 'The re-entry vehicle will splash down between 03:05 and 03:08 GMT'. Thus they seem to equivocate on the term, by shifting between the terms 'prediction' and the more demanding-sounding 'point prediction'.

Bernstein et al. (2000: 54) say that, '[s]cenarios make contingent claims rather than point predictions'. Thus point predictions may be contrasted with more general predictions, and Bernstein et al. contrast them with 'contingent causal claims'. What does this term mean? It would seem that all predictions are contingent statements. What sort of statement is *not* contingent? The usual contrast is to 'necessary' statements. So any predictions that are not 'contingent statements' would be about alleged necessary future events, i.e., 'inevitabilities'. But many, if not most, IR theorists who hazard predictions avoid the concept of 'inevitability'. Of course there are some highly deterministic theorists who talk of 'historical inevitability'. But this is not required of theoretical prediction in IR. There are many reputable theorists who develop covering laws and theories and offer predictions, but avoid insisting that they are offering 'necessary' truths.[26]

Foreknowledge and revisability

The second point Bernstein et al. make in the passage above is that theory-based prediction entails claims of 'knowing the future', while scenario analysis allows for uncertainty and thus works with the unfolding of new information as the policy-interactions proceed. But with respect to this 'foreknowledge' problem, it is possible to see that it is inaccurate, particularly in view of the fact that the predictions of IR theories are probabilistic and contingent. So the criticism of Bernstein et al. misses its target. There is no such sharp difference between the two methods of IR theory and scenario analysis. The outcomes are not known until after the fact in either the scenario-based or theory-based methods, especially since theory-based methods in IR involve probabilistic predictions, which preclude, at least in most cases, 'knowing' in advance what the outcome will be.

Plausible assumptions and 'probability'

The third point Bernstein et al. make is that scenario analysis allows the use of the most plausible assumptions, which can be revised as time passes and new information arises, while it avoids unwarranted theoretical prediction. But, given a particular set of policy problems, the policy-maker must choose (a) which scenarios to write and (b) what consequences follow from what conditions, that is, which claims qualify as the 'most plausible'. Both a and b are dependent upon the causal connections that are viewed as the most acceptable or plausible; which connections are considered plausible depends on what theories the policy-maker finds acceptable. Bernstein et al. say that actors 'evaluate decisions against the most plausible scenarios in the current set' (2000: 57). How does the decision-maker or analyst answer the question 'Which scenarios are the most plausible?' The answer again is IR theory, which systematises a mass of data and yields probabilistic associations. How else can the analyst distinguish an implausible from a plausible scenario? Bernstein et al. add that policy-makers then evaluate 'the likelihood of these scenarios as their strategy unfolds' (2000: 57). What sort of basis can there be for such an evaluation? Again, the evaluation is a result of a similar process after recalculating probabilities using any further accumulation of data.

On this point Bernstein et al. say that '[t]he foundation for scenarios is made up of provisional assumptions and causal claims' (2000: 54). But why do analysts select one set of provisional assumptions rather than the thousands of other possible sets? And what basis is there for the causal claims that are part of the playing out of the scenario? Again, the answer can only be derived from generalisations based on the available empirical data, which are systematised into causal relationships by means of theory. While it may be true that a term whose use is as widespread as 'causality' has more than one meaning, Chapter 4 argued that any adequate analysis of the term reveals that observed generalisations form part of the basis for causal claims.

The use by Bernstein et al. of NATO's decision to bomb Serbia can be seen as a case in point. Suppose that in 1999, as NATO was moving closer to a decision to use force, experts in scenario-construction approached defence planners at SHAPE, the Pentagon or defence ministries in Western Europe and offered three scenarios. First, President Milosevic orders stepped-up ethnic cleansing to create a *fait accompli* before NATO intervention can begin; second, he demands a UN vote condemning any alliance use of force without UN approval; and third, to shame the leaders of member states and as a way to protest against even a consideration by NATO of forceful intervention in the internal affairs of a sovereign state, he orders half of his Yugoslav regulars and paramilitary units to commit suicide outside their barracks. NATO defence ministers will not request any analysis or discussion of the third scenario because all scenarios, as Bernstein et al. point out, begin with a set of initial or boundary conditions which they call 'provisional assumptions'. Different sets of such assumptions are 'most plausible', 'less plausible', 'still less plausible', etc. (Bernstein et al. 2000: 57). What this means is simply that they have higher or lower probabilities, given one's background knowledge and beliefs. Why do the defence ministers reject any further discussion of the third scenario? They do so because its assumptions have near-zero probability (and no high disutility or potential damage to Western national interests if it were to come about).

On what basis do analysts or policy-makers assert that there is a near-zero probability? On the basis of a theory that systematises the wealth of empirical data, including the fact that no sovereign head of government has ordered his/her forces to commit mass suicide before an engagement and that there is nothing particularly unique about the Kosovo conflict or Slobodan Milosovic that would lead them to think that well-established and well-understood patterns of non-suicides in the past would be violated (and so to assign any unusual probability values to the propositions). These analysts' and policy-makers' theories, whether clearly and consciously articulated or not, do not permit any cause-and-effect reasoning that would connect the circumstances of the Kosovo crisis in 1999 to Milosevic's issuance of the suicide order.

Compatibility of scenarios and probabilistic predictions

Point (iv) above, that the two methods are fundamentally at odds, does not seem warranted. The two methods are not fundamentally opposed to one another in that the scenario method is an effective way to help conceptualise various possible outcomes. The method of scenario-construction Bernstein et al. endorse is, in fact, compatible with probabilistic prediction. Especially since prediction in the social sciences and much of the natural sciences is probabilistic, the scenario method can best be seen as an extension of probabilistic prediction (or of the application of rational choice theory that includes probability estimates). Each scenario is a detailed outcome of a decision tree or cell of a decision matrix. The scenarios themselves are chosen on the basis of beliefs about probability assignments (and utility estimates), which is why there is no scenario for a French invasion of China or President Milosevic pre-emptively ordering his troops to commit suicide outside their barracks.

Probability estimates are a crucial part of the policy-planning process, even if there are other factors to be considered, such as (dis)utility, moral, legal, etc. For example, low probability/high disutility scenarios involving nuclear strikes are gamed by defence planners, while low probability/low disutility scenarios, like the Serbian suicide protest, are not. The traditional methods of estimating probabilities, based on the best available theories, which are combined with utility estimates to analyse policy choices is parallel to the 'wild-card' scenario analysis Bernstein et al. recommend. They say that there 'are conceivable, if low probability, events or actions that might undermine or modify radically the chains of logic or narrative plot lines of scenarios' (Bernstein et al. 2000: 58). These are analysed by means of scenarios that are highly detailed game theory matrices. A breakdown of a typical scenario reveals the set of alternative choices that standard game theory matrices offer.

There is clearly room for 'prediction' in IR, when the term is understood along the lines of the definition on p. 8, since it is much less demanding a notion than that of 'point prediction'. The latter would have the form statements like those above regarding a Sino-French war in 2013 or the splash down of the re-entry vehicle, in contrast to a statement that meets the definition of 'prediction' *simpliciter* on p. 8, like 'Wars are more likely in multipolar systems than in bipolar systems' or 'China will not be unilaterally invaded by France in the next decade'. Bernstein et al. *do* allow prediction, at least given the definition on p. 8.

Bernstein et al., prediction and scenarios

This section has sought to point out flaws in the arguments of Bernstein et al. regarding the definition of IR terms, the uncertain relationships between underlying and immediate causes, problems of social learning, the single-case problem and the argument of Doran drawn from non-linearities of social behaviour. Bernstein et al. criticise the 'received view' of the natural sciences, which takes a positivist-empiricist interpretation of Newtonian mechanics as the paradigm of scientific knowledge and the attempt to model the social sciences on it. They proceed to offer the alternative of scenario analysis. Bernstein et al. correctly argue that scenarios are a helpful way of planning policy. They argue erroneously that traditional theory-based prediction in IR is inconsistent with scenario analysis.

In their attack on the Newtonian analogy, Bernstein et al. seem to portray Newtonian mechanics and its foundations as perfectly well-understood by all physicists and philosophers, in contrast to social science theory and its foundations, which they portray as unsettled and subject to doubt. While they are right in holding that over-emphasis of the analogy with classical physics can be troublesome, they err in the opposite direction by overstating some of the differences between Newtonian mechanics and social science theory. Physical theory has much more unanimity than IR, to say the least, on the matter of the best set of laws and the most powerful theory. However, there is not so much more unanimity on the philosophical foundational questions, e.g., regarding the meanings of the laws and theories. As in IR, the foundations of physics are very much subject to debate among philosophers of science, e.g., in the post-Newton world, between conventionalists, instrumentalists, operationalists, constructive empiricists, logical positivists, scientific realists and others. There is far more of a parallel between the foundational debates of physical science and social science than Bernstein et al. recognise. They thus pose a false dichotomy between the former as stable and settled and the latter as shambolic.

The attempt by Bernstein et al. to distinguish theories of physical science from those of social science fails on at least three grounds. First, physicists have to agree on theories in order to agree on the meanings of terms, which is inconsistent with the distinction between theoretical terms that they draw (though they do seem to be aware of this in some of the specific comments they make). Second, the assessment of truth-values of theories is not so radically different between natural and social sciences, at least on the best accounts of natural science. And third, idealisations are not only part of the social sciences but, as Cartwright (1983) has shown, and contrary to the assumption of Bernstein et al., they are crucial to classical theories of natural science in the form of frictionless pistons in cylinders, ideal gas laws, etc.

Bernstein et al. minimise the accomplishments of IR theory by asserting that the ‘truths’ that IR theorists have discovered ‘are close to trivial’ (Bernstein et al. 2000: 44). However, it is clear that when some apparently common-sensical platitudes turn out, under scrutiny, to be false, e.g., ‘States that have the power and resources to win wars are usually the initiators’ (Arquilla 1992, Arreguín-Toft 2001, Paul 1994). It takes good scholarly analysis and effort to discover whether truisms are true or false, and the discovery of their truth value is genuine progress. There are scarcely any ‘trivial truths’ that do not require empirical or theoretical support in IR. Moreover, some initially contested DP claims have, over time, become quite well established (see Chapter 6).

The more Bernstein et al. can explore specific characteristics to show that IR is like evolutionary biology, which is non-predictive, and the less like classical mechanics, which is predictive, the more they motivate the conclusion that IR is non-predictive. In general, the greater the number of well-known characteristics A is shown to share with B and the fewer with C, the more reasonable it is to conclude that the A will share unknown characteristics with B and not with C. This is good inductive reasoning. But if the argument is to be persuasive, it must have both valid form and true premises. This chapter has sought to cast doubt on the premises, especially the claim that IR is much more like evolutionary biology than classical mechanics.

Popper's critique of social prophecy

One might include Popper in the list of the philosophical critics of prediction in the social sciences. Popper argues vigorously in *The Open Society and Its Enemies* (1971a, 1971b) and *The Poverty of Historicism* (1945) against an expansive notion of 'prediction' in the social sciences. (See also Popper 1948). However, there is a crucial difference between the position of Popper and those of the prediction sceptics just discussed, namely, Popper's sensitivity to the different notions of 'prediction'. Bohman, Doran and Little all use a single, rather grand notion of 'prediction' and ignore more modest and realistic notions. Popper attacks the grandest notions but, in contrast to others, clearly distinguishes several notions of prediction – e.g., prophecy versus technological prediction (1971a: 157, 1945: 43–5), long-term versus short-term prediction (1945: 36–8), prognosis versus prediction (1945: 133). This section seeks to show that Popper was a kind of naturalist, that he did not reject the justifiability all possible notions of 'prediction' in the social sciences and that he endorsed a notion roughly parallel to the sort of prediction this book defends, defined on p. 8.

As noted in Chapter 1, Popper focuses much of his attention in *The Open Society and Its Enemies* (1971a, 1971b) and *The Poverty of Historicism* (1945) to attacking historicism, the deterministic view that history moves inevitably and inexorably toward a specific state of affairs. Popper also considers and rejects an alternative position, which he calls 'utopian social engineering' (1971a: 157–68, 1945: 42–5). Popper criticises the latter view because, like historicism, it requires the acceptance of deterministic connections between one social state of affairs and another. Utopian engineering holds that implementation of policy A will inevitably bring about condition B, which will eventually bring about condition Z, which is the final goal of human society.

Popper's criticisms of historicism reveal his own support for a form of naturalism and social prediction. He says that some elements of historicism are naturalistic (scientistic) and some are not; e.g., historicists regard social experimentation as useless (1945: 35). Popper suggests that he accepts a form of naturalism right from the first sentence of the first chapter of *The Poverty of Historicism*, which specifically contrasts historicism with 'methodological naturalism' (1945: 4). But he seems opposed to methodological naturalism (1945) – or scientism (1971a, 1971b) – when he describes it as 'the tendency to ape, in the field of social science, what are supposed to be the methods of the natural sciences' (1971a: 286 n. 4) – especially by his use of the pejorative verb 'to ape'. But in fact Popper is endorsing naturalism here, as he immediately goes on to declare that 'the methods of the social sciences are, to a very considerable extent, the same as those of the natural sciences' (1971a: 286 n. 4). Popper does support the idea of social laws, at least of a very general sort and offers an example.[27]

Popper also supports a notion of 'prediction' in the social sciences. But it is a notion much weaker than historicist writers assume. What Popper rejects is deterministic point prediction, like the astronomer's prediction of an eclipse (1971a: 186 n. 4). Social phenomena, like revolutions, cannot be predicted in a strictly parallel way. But Popper's support of a weaker notion of 'prediction' is evidenced by his endorsement of the possibility of 'piecemeal engineering' (in contrast to 'utopian social engineering').

Among historicism's naturalistic elements is 'the importance of prediction as one of the tasks of science' (1945: 12). Popper's support of a limited form of social prediction is clear in the next sentence, as he adds parenthetically, 'In this respect, I quite agree with it, even though I do *not* believe that *historical prophecy* is one of the tasks of the social sciences' (1945: 12). Elsewhere, in a very similar passage, Popper says that, according to historicism, 'success in sociology would likewise consist, basically, in the corroboration of predictions. It follows that certain methods – prediction with the help of laws, and the testing of laws by observation – must be common to physics and sociology' (1945: 36). But Popper immediately adds, 'I fully agree with this view, in spite of the fact that I consider it one of the basic assumptions of historicism.'

The problem with historicism, according to Popper, is not that it seeks to predict in the sense defined in p. 8 above. Rather, it goes awry by failing to acknowledge various qualifications that that definition of 'prediction' includes. Most significant is that historicism offers 'social prophecy' based on trends that historicism regards as deterministic and that it does *not* regard as contingent. Popper says, 'the central mistake of historicism [is that] its *"laws of development" turn out to be absolute trends*; trends, which, like laws, do not depend on initial conditions, and which carry us irresistibly in a certain direction in the future' (1945: 128). Popper holds, rather, that social laws and the predictions based upon them, are contingent upon initial conditions and are non-deterministic.

Thus Popper sees prediction as an appropriate aim of the social sciences, as long as the term is properly understood – not as deterministic point prediction, inevitability or even long-term trends. But a more circumscribed notion of 'prediction' is proper within the social sciences. While Popper's arguments are considerably different from those advanced here, his conclusions rejecting deterministic point prediction in the social sciences and supporting a more limited (probabilistic) notion are consistent with the conclusions defended in this study.

Conventionalism, causation and prediction

Does the method of Duhemian conventionalism described in Chapter 4 allow for prediction in the way that does not seem possible on the meta-theories discussed above? Duhem's own work is concerned with the physical sciences and, according to his account, prediction is a fundamental element of scientific method and the basis for evaluating theories. Duhem advocated a type of falsificationist methodology (though Popper regarded it as too verificationist, e.g., 1968: 78–81), according to which theory-testing is possible only by deriving consequences about expected outcomes and conducting experimental tests to determine if the expectations are met.

Duhem identifies flaws in the traditional understanding of scientific theory-testing, along the lines discussed in Chapter 4. For Duhem, there are no crucial experiments. Given two theories that produce two different hypotheses, H1 and H2, if, under conditions K, H1 → A (that is, under conditions K, if hypothesis H1 is true, event A will be observable) and H2 → not-A, one may set up an experiment to determine whether A or not-A is true (or obtains under conditions K). However, if A is shown to be true, then, contrary to the standard approach, one is not entitled to conclude that H1 is true and H2 is false. There may be another theory that produces H3, where H3 → A. Since A follows from both H1 and H3, one cannot then determine, without conducting further tests, whether H1 or H3 or some as-yet unthought-of hypothesis H4 is true. So one may not conclude that H1 is true, because there may be an as-yet unformulated theory that is true and performs better.

Similarly, one may not conclude that H2 is false when A is shown to obtain in the experiment. What is clear is that the conjunction of the theory, H2 and all of the auxiliary assumptions that one must make in order to derive the consequences and relate observed events to the hypothesis, has a falsehood among its conjuncts. But it may be that H2 itself is not the culprit. So any given theory may be saved in the face of apparent anomalies by rejecting an assumption or hypothesis to create consistency with the observed evidence. (Poincaré agrees that theories are hard to refute but draws this conclusion for very different reasons.)

This does not, however, mean that Duhem believes that theories can never be refuted by empirical evidence. Duhem believes that an anomaly may require the investigator to reject the assumption that the experimental apparatus is in good working order. But anomaly after anomaly will not reasonably permit the blame to fall on the assumption of a failed apparatus. After repeated anomalous results, the collective weight of evidence against the theory will require that the investigator 'of good sense' and the community of investigators give up the theory in favour of a rival.[28]

Duhem holds that empirical evidence can force the rejection of theories and, as such, he is a thorough-going fallibilist. As noted in Chapter 4, Poincaré holds that theories may be maintained in the face of anomalies because laws have the status of definitions. A freely falling body accelerates at a specified rate. If experiment finds a body that does not obey this law, it has violated the definition and cannot thus be properly regarded as a 'freely falling body'. Nation-states act to preserve their survival; if an entity does not act to preserve its survival, on Poincaré's analysis of theories and laws, it turns out not to have been a nation-state.

Duhem rejects Poincaré's approach to 'laws' and argues instead that experiments can prove and disprove theories, though not singly, i.e., in the form of the individual, decisive and crucial test. The way experiments serve to indicate which is the superior and which is the inferior theory is by means of drawing out each theory's predictions and comparing the predicted results with the actual results *over a period of time* using a range of different tests. So despite the fact that

Duhem sees various limitations on the traditional understanding of scientific testing, he clearly believes that testing is possible – and indeed indispensable – and that prediction is possible and necessary.

Is it possible to consider covering laws as part of social-scientific theories? Such an extension of the traditional natural-science method requires the application of covering laws to the world of probability. But such an application should not be problematic, and is indeed necessary in a variety of areas of natural science, from quantum physics to genetics. Rather, a clear understanding of the differences between Duhemianism and more well-known positivist meta-theories shows that Duhemianism fits more neatly with social science theory-testing and hypothesis-testing.

For example, one might formulate a probabilistic law like (HM) *there is a 0.75 chance that a national leader in a crisis will sacrifice crucial intelligence rather than tangible military forces*. One then selects as the crucial test the case of Pearl Harbor. President Roosevelt had to choose between protecting tangible military forces or intelligence assets. The investigator notes that the outcome chosen by Roosevelt was contrary to the expectation of HM. However, a defender of HM could point out that the law only states that the choice of physical forces over intelligence is expected 75 per cent of the time and the Pearl Harbor counter-example is one of the 25 per cent of cases where the opposite choice is made. Duhem's account of scientific theory-testing rejects crucial experiments, which is precisely what an account of the social science disciplines, with their probabilistic laws, must do.

The shift from Newtonian deterministic laws to social-scientific probabilistic laws is not a problem on the Duhemian account, since theories are displaced by rivals only by the accumulated weight of counter-evidence gathered over a period of time involving many cases or instances. Thus the Duhemian understanding of theory-testing and the development of a scientific discipline, while fallibilist and falsificationist, is not tied to any notion of simple falsificationism and does not run counter to the constraints that attend probabilistic generalisations.

Conclusion

Over the past century IR developed as a discipline with the primary goal of providing a basis for changing the world for the better; i.e., as a basis for the formulation of policy. Despite the fact that most critical theorists, reflectivists and anti- and post-positivists passionately wish to use the study of IR for that purpose, the meta-theories they adopt, with their attacks on prediction, undermine the project of using the study of IR to change the world.

In IR and other social sciences, policy-makers must predict and, contrary to post-positivism, if there are no sounder generalisations available, then phenomenal regularities must be enough to go on, at least some of the time. Even for those who accept Bohman's HT arguments against deterministic explanation in the social sciences, and who thus see explanation as perspectival, incomplete and circular, it is possible to ground a notion of 'prediction' that is capable of satisfying enabling conditions (at least closely enough) to allow policy-makers to use social theory generalisations as a foundation for predictions and thus for policies. The generalisations are neither perfectly reliable nor deterministic. They are probabilistic and offer the decision-maker imperfect guarantees about the future. Still, in many epistemic circumstances, they provide the decision-maker with much greater assurances that the chosen policy will lead to the desired result than he or she would have if policies were selected randomly.

Bohman overlooks the need for social prediction even more than the other authors discussed. Little overlooks it, but at least makes reference to 'probabilistic prediction' in some of his writings. Doran begins with an argument for prediction-scepticism but then adds several qualifications, which would seem to open up some room for prediction or forecasting. However those remarks are surrounded by comments that undercut the qualifications and which thereby restore full-blown prediction-scepticism. Bernstein et al. offer a range of criticisms that fail to target the most common sorts of predictions of policy-makers and they offer an alternative that makes use of the sort of theory-based prediction they claim to have rejected. This chapter has thus sought to show that the arguments against prediction offered by each author are flawed and that the sound elements of the foundational positions sketched out by the various authors (especially Bohman and Bernstein et al.) can consistently be brought into line with some notion of 'prediction', when that notion is founded on probabilistic rather than deterministic generalisations.

Bernstein et al. attempt to discredit 'prediction' by arguing that IR is much more similar to evolutionary theory than to physical sciences like classical mechanics. Is IR very like classical mechanics or evolutionary biology? It shares many features with both but also has many dissimilarities to both and consequently is 'very like' neither. A major part of the strategy of the critique of Bernstein et al. has been to show that a further probing of the character of physical science reveals that the dissimilarities that Bernstein et al. claim do not hold. This is not to say that an unrestricted naturalism is justified. Far from it. Comprehensive theories like those of the physical sciences are not likely to emerge in IR. Nevertheless, theoretical and scientific-style investigation in IR has great value and holds out the possibility, at least within tightly circumscribed domains, to achieve natural-science-like consensus and well-founded prediction. Indeed, prediction is necessary for good policy-making, even though there are limitations due to hermeneutic interpretation, lack of governing regularities and non-linearities. These considerations lead to the conclusion that there are limitations on the types of predictions one might propose and the confidence that should be displayed in them but not to conclude that policy-makers should avoid prediction. While prediction is necessary for policy-making, prediction alone is not sufficient, since normative considerations must always be addressed. Probabilistic predictions may inform one of things like 'socialist states go to war with non-socialist states less often than democratic states go to war with non-democracies'. But normative analysis is clearly required in order to determine whether this is a good or bad thing and what policy initiatives should be pursued.

Even though it is not the case that a very strong version of naturalism follows from the preceding argument, nevertheless, the relationship of IR to physical and biological science is much more complex than Bernstein et al. seem to acknowledge. And the problem of non-linearities does not completely undermine predictiveness, as Bernstein et al. and Doran seem to think. IR theory, or at least a good portion of it, does not attempt to generate point predictions like those of the physical sciences. Part of the complex mix of physical and biological sciences is at least a degree of predictiveness of IR theory with specific, identifiable limitations. This characteristic is essential, as at least Bernstein et al. recognise, if IR theory is to have any hope of changing the world.

This study adopts a pragmatic orientation, which is clear from the endorsement of aspects of the epistemology of Peirce in Chapter 3 and of the conventionalism of Duhem in Chapter 4. The pragmatism necessary for the practising foreign policy-maker can be seen in the answer to Little's charges against the predictiveness of social science theory considered above on pp. 138–42, which draws on his account of 'causation' outlined in the last chapter. Little argues that there are too many variables, the chains of reasoning are long and, since the connections are probabilistic, the longer the chain, the lower the probability of the prediction.

There are, though, applications of IR theory where the chains are not long. When the question at issue is epistemically local, it will be possible to make high-confidence (but obviously still fallible) predictions. There are more global or distant predictions that will be tempting. But the guidelines discussed above will lead to a lower level of confidence in predictions in those cases. Kyburg's (1961) argument against the lottery paradox can be taken as a model. In any case, there will be costs associated with accepting policies based on those low-confidence predictions. However, for the practising foreign policy-maker, those costs must be weighed against the costs of suspending judgement and taking no active decision.

If meta-theoretical and philosophical arguments show conclusively that prediction in IR and the social sciences is impossible, then such conclusions must be taken seriously. If they show that there is some rational way to base policy without prediction, then that method should be explored. But the present chapter has sought to show that the current state of the debate, using the most tenable versions of each of the three sources of anti-predictive argument, all fail to sustain anti-predictive conclusions and can be shown to be consistent with a notion of 'prediction' that provides the necessary foundation for inferences about connections between present policies and future results. It has been argued that conventionalism generates an account of 'prediction' and the quasi-Duhemian version defended here endorses causal connections. It is possible to formulate a distinctive version of conventionalism, applicable to the social sciences, according to which more broadly drawn statements (and possibly models), qualified with a degree of confidence, can rationally be formulated about the future.

6 Explaining agreement and disagreement in the natural sciences and social sciences

EINSTEIN: I'm supposed to meet her at six o'clock at the Bar Rouge.
FREDDY: This is not the Bar Rouge. It's the Lapin Agile.
EINSTEIN: No difference.
FREDDY: No difference?
EINSTEIN: You see, I'm a theorist and the way I see it is that there is just as much chance of her wandering in here accidentally as there is of her wandering into the Bar Rouge on purpose. So where I wait for her is of no importance. It is of no importance where I tell her I will be. And the least of all, it's not important what time I am to meet her.
FREDDY: Unless …
EINSTEIN: Unless what?
FREDDY: Unless you really want to meet her.
EINSTEIN: I don't follow.
FREDDY: If you really want to meet her, you'll go to the Bar Rouge at the time you told her.
EINSTEIN: You're forgetting one thing.
FREDDY: What's that?
EINSTEIN: She thinks like I do.

Steve Martin, *Picasso at the Lapin Agile* Act I

Theorists of IR, like those in other areas of social science, have had difficulty reconciling the contending research traditions, for example those of realism and liberalism.[1] Unlike the social sciences, the physical sciences are generally regarded as having cumulation of knowledge and approach-to-consensus in the history of the discipline.[2] Cumulation and consensus in IR have been slow in coming. Three-quarters of a century ago, John Hobson felt the need 'to afford some explanation of the slowness of these sciences in producing any considerable body of larger truths, in the shape of generally accepted laws and principles' (1926: 7). The need has only increased since then. The adoption of the thesis of CS raises several questions. Why does there continue to be disagreement in some fields of enquiry but not in others? Why is it that the physical sciences display cumulation and approach-to-consensus over time, while IR, in contrast, with few exceptions, does not? And why are there those particular exceptions?

The fact that CS raises these questions suggests a reconsideration of one of the criteria of theory choice discussed in Chapter 3: Lakatos's criterion of 'new facts'. Chapters 3–6 argue that CC theory provides better answers to important meta-theoretical questions of 'law', 'causation' and 'prediction', than more popular and well-known accounts. Lakatos's criterion of 'new facts' states that a good theory should solve problems other than those it was intended to solve, and it should point the investigator to previously unknown facts and questions and should offer answers to those questions. As noted, CS and CC raise the question of why there is sometimes consensus and sometimes dissensus in various fields of empirical enquiry. This chapter argues that CS and CC also offer the best answers to these questions.

Three more specific questions then stand in need of answers:

i) Why have proponents of different traditions in IR been unable to agree on the superiority of one theory or approach?
ii) Why has this happened in IR while some other empirical sciences, such as physics, do not have comparable diffusion? And
iii) why have some areas of IR, like research on the DP hypotheses, been able to approach consensus on some important theoretical matters?

Thus, this chapter seeks to show both that CC, as a descriptive and explanatory account of IR, is a progressive one in Lakatos's sense (in the first four sections) and that the discipline of IR is capable, as CC emphasises, of specific research programmes that are progressive (in the last four sections). The latter is done using the example of studies of the DP hypotheses. They are shown to have a hard core, which has remained intact throughout twenty-five years of intensive empirical and theoretical scrutiny and are shown to exhibit cumulation and approach-to-consensus in part by means of a social science analogue of the physicist's measure-stipulation.

Lakatos's criterion

Lakatos's methodology

Lakatos offers the most widely endorsed methodology of IR. While Popper and Kuhn are perhaps cited more frequently, they are cited both favourably and critically. Since Lakatos commands so much loyalty among methodologists of IR (see below), it is worth considering CC within the criteria he adumbrates. Lakatos advances an account of scientific method, inspired by Popper, which he terms 'sophisticated falsificationism' (1978: 37–47). The philosophy of science has produced philosophically defensible criteria of theory choice derived from deduction, as Popper prefers, and from other sources. Lakatos, Duhem and others offer examples of such criteria. Duhem's are, in part, incorporated in the CC account above. Lakatos defends 'progressiveness' and 'prediction of novel facts', among others. He says that an emendation of an existing theory is 'progressive' if it 'leads to the actual discovery of some new fact' (1978: 34). In the present study the adoption of the CC account of IR theory leads to the discovery of reasons for certain observable patterns in the history of IR theory, much as Lakatos was interested in discovering the reasons for the patterns in the history of physics.

One might consider the range of collected observations in a given domain and then construct three distinct sets of laws and explanations to fit them. How would these three be evaluated? One measure would result from the distinction Lakatos draws between 'progressive' and 'degenerating' research programmes. The difference between theories or research programmes exhibiting varying levels of progressiveness becomes evident when repeated *ad hoc* adjustments are needed to accommodate new observations; repeated *ad hoc* adjustments are a sign of a degenerating research programme. Progressive research programmes not only need fewer *ad hoc* adjustments but they actually point the way to new facts and explain phenomena that the theory was not specifically formulated to explain. Lakatos argues for the superiority of progressive research programmes over degenerating ones.

Popular acceptance in IR of the Lakatosian account

A wide variety of contemporary IR theorists endorse some version of Lakatos's criterion of the 'progressiveness' of scientific research programmes or his notion of 'sophisticated falsificationism'. A few references to the most influential figures in neorealism, liberal institutionalism and reflectivism should provide a sense of the wide acceptance of Lakatos's methodological ideas.

Most of the prominent current IR theorists endorse a Lakatosian approach to theory-testing and appraisal, though not all of them explicitly discuss each of the major components of Lakatos's philosophy of science. Their adoption of Lakatos's views is sometimes suggested by their use of distinctly Lakatosian terminology, viewing theories in terms of 'research program(me)s', with 'hard core' and 'protective belts' of propositions, even if they do not specifically cite Lakatos.

Among the neorealists, Waltz (1997) directly engages Lakatosian criteria to evaluate neorealism. Waltz says, 'Because of the interdependence of theory and fact, the construction and testing of theories is a more problematic task than most political scientists have thought. Understanding this, Lakatos rejected "dogmatic falsification" in favor of judging theories by the fruitfulness of the research programs they may spawn' (Waltz 1997: 914). Waltz argues that the attacks on neorealism based on the criteria of 'progressiveness' and 'fruitfulness' are erroneous, not because Lakatos is mistaken in advocating such criteria, but rather because the criticisms (by Vasquez 1997, 1998, 2002 and others) mischaracterise both neorealism and the Lakatosian criterion. Other neorealists are considered below.

On the anti-realist side, Keohane, Russett and Vasquez all appear, at least at some point, to appeal to Lakatosian criteria of 'prediction of novel facts' or 'fruitfulness'. Keohane (1989: 43–4, 52, 55–6, 59) discusses the inability of neorealism to satisfy Lakatos's criterion of fruitfulness (see also Vasquez 1998: 243). Keohane and his colleagues single out the criterion of 'fruitfulness' or 'prediction of new facts' in their first statement of what 'demanding tests' must be passed for scientific hypotheses to be considered certain. They say, '[a]t minimum it must be consistent with our knowledge of the world; at best, it should predict what Imre Lakatos (1970) refers to as "new facts", that is, those formerly unobserved' (King, Keohane and Verba 1994: 12).

Russett discusses the 'empirical assumptions at the hard core of the hegemonic stability research program' in his critique of the assessments of America's declining hegemony (1985: 230–1). Russett and Oneal (2001) seem to endorse the criterion of prediction of new facts when they say that truth 'shines unexpected light into many corners' (2001: 44).

Vasquez uses the Lakatosian criterion as a centrepiece of his critique of neorealism in his contribution to the APSR symposium (Vasquez 1997) and in the revised edition of *The Power of Power Politics* (Vasquez 1998; see also Vasquez 2002). He considers various sets of criteria of adequacy for scientific theories or research programmes, which should be applicable to theories of IR and endorses: accuracy, falsifiability, explanatory power, consistency with what is known in other fields, parsimony or elegance, and the explicitly Lakatosian 'progressiveness of research programmes' (Vasquez 1998: 230). Vasquez (1998: 230) says, 'a criterion that is of great relevance to the inter-paradigm debate is research programs must be progressive rather than degenerative. This is the key criterion used by Lakatos.' Vasquez attempts to show, 'contrary to widespread belief, the theoretical fertility that realism has exhibited in the last twenty years or so is actually an indicator of the degenerating nature of its research program' (1998: 242). Vasquez acknowledges Lakatos's criterion and connects theoretical fruitfulness with the progressive or degenerating nature of theories. It is also interesting for present purposes to note that he asserts that there is any sort of 'widespread belief' about Lakatos's criterion and political realism. Vasquez cites Hollis and Smith (1991: 66) and Wayman and Diehl (1994: 262) as endorsing the Lakatosian criterion.

Vasquez's use of Lakatos's criteria of 'fruitfulness' and 'progressiveness' in his attack on the neorealism of Waltz, Schweller, Walt, Elman and Elman and Snyder and Christensen forces those authors to engage in the enterprise of theory appraisal in light of Lakatos's widely accepted criteria of theory choice. Walt alone rejects the Lakatosian criteria (Walt 1997: 931–2). Waltz (1997) and Elman and Elman (1997) specifically endorse at least a part of Lakatos's methodology. Both replies to Vasquez contend that Vasquez's mistakes are misunderstandings of Lakatos's criterion and mis-specifying neorealism. Snyder and Christensen (1997) similarly endorse Lakatos's criterion, stating, 'John Vasquez is right to insist that students of international politics should justify their theories in terms of Imre Lakatos's (1970) criterion for distinguishing progressive research programs from degenerative ones' (1997: 919). Schweller (1997), like many contemporary IR theorists, accepts Lakatosian methodology without any critical evaluation.

Jervis, also generally regarded as a realist, and certainly one of the most important contemporary theorists, cites Lakatos several times in his classic *Perception and Misperception in International Politics* (1976) on the general issues of the difficulty of objective criteria for theory choice. He does not probe the specifics of Lakatos's criteria or arguments. But he cites Lakatos and Kuhn in developing his arguments against prediction (Jervis 1976: 414–15), discussed above in Chapter 5.

Postmodern and constructivist writers make use of the Lakatosian criteria, at least when evaluating naturalist theories. Ashley addresses this aspect of the criterion directly, noting that some neorealists see their theory as part of a 'progressive scientific' discipline (Ashley 1986: 260). Ashley uses the Lakatosian understanding of 'scientific research program' to relate neorealism's commitment to 'statism' as its 'hard core' and 'protective belt of auxiliary hypotheses', which are derived from an eclectic array of sources (1986: 275–6). The constructivist Wendt says the only criterion of theory choice he endorses is the Popperian and Lakatosian criterion of falsifiability (1999: 89).

> All scientific theories must meet the minimum criterion of being in principle falsifiable on the basis of publicly available evidence, and social scientists should approach their knowledge claims with that in mind. Beyond this, however, we should be tolerant of the different standards of inference needed to do research in different areas.

And he adopts a Lakatosian conception of the character of theories when he refers to the 'hard core of research programmes' (1999: 28).

Finally, it is noteworthy that even an avowed opponent of philosophical musing for its own sake, like Valerie Hudson – a self-described 'IR mechanic' (Hudson 2001) – is compelled to endorse Lakatos's methodology. She says, referring to the formulation and testing of foreign-policy analyses, 'If meta-theory cannot be translated into something that would have practical application to this type of endeavor, I admit to having little use for it' (2001: 1) but later adds, 'I am an unrepentant Lakatosian' (2001: 4–5).

The Lakatosian criterion of 'new facts'

According to Lakatos, theories that are found to have anomalies are typically revised in light of such counter-evidence and evolve into new theories. The original theory should be regarded as evolving into the new and more adequate version rather than as having been falsified and needing to be discarded. Had the original theory simply been viewed as falsified and then discarded, the result would have led to the rejection of many of the most successful and widely accepted theories in the history of science. So for Lakatos the unit of appraisal is a ‘series of theories’. When a shift to a new series of theories occurs it is because the new set is superior in that it meets several crucial criteria. One is that the new theory includes the (unrefuted) portion – or, as Popper and Lakatos conceive it, ‘content’ – of the previously accepted rival theory set. Second, the new theory adds additional *content* beyond the previous theory. The new content must be corroborated, which is evidenced by the new theory ‘predicting’ or ‘leading to the discovery’ of what Lakatos calls ‘novel facts’.

These facts or observed phenomena may be predicted to occur in the future. Lakatos often uses the phrase ‘prediction of novel facts’ with reference to this criterion. But they may also have occurred in the past; Lakatos says explicitly (1978: 32 n. 3) that ‘postdiction’ (also termed ‘retrodiction’) is included in the ‘wide sense’ in which he uses the term ‘prediction’. The phenomena may have occurred in the past and also have been previously observed, as is clear from Lakatos’s example of the differentials in astronomical measurements taken during the day versus those taken at night. The possibility of carrying out such measurements was always there. So for Lakatos, a progressive research programme takes older theories and replaces them with new ones that both add new content and predict new facts.

The CC account of social science theory offered here meets the Lakatosian criterion of ‘prediction of novel facts’, i.e., of explaining observed patterns that had not heretofore been adequately explained. In this case the patterns that need to be explained are those of agreement and disagreement or consensus and dissensus in empirical (natural and social) sciences. The history of natural science, and especially of physical science, is one of approach-to-consensus over the long term. In the physical sciences many theories debated three hundred years ago are universally rejected today and there is convergence, even if not unanimity, on many of the most general theoretical principles. In IR the most general theoretical principles, about the possibility of joint gains, cooperation, and the possibilities for international organisations, that were debated three hundred years ago are still debated (though there has been some evolution of their formulations).[3] On a more micro-level examination of the fields, one finds that there are areas of convergence within IR, such as the enquiry into the DP hypotheses. The conventionality principle offers a way to explain of these patterns, which is not possible on many of the rival accounts based on SR, the IP thesis, the RU thesis and HT.

The inescapability of conventionality

This book has built a number of conclusions on the premise that physical theory is inescapably conventional. This claim can be substantiated by looking at the recent history of physical theory. Euclidean geometry was regarded as the geometry of physical space for centuries before Newton. The extraordinary success of Newton’s physical and astronomical theory reconfirmed this view and left little doubt about its truth. A century later Kant argued not only that Newton was right to incorporate Euclidean principles into physical theory but that those principles are synthetic *a priori* truths. For Kant space, like time, is a form of sensible intuition: one can have no experience of the physical world that does not conform to the geometry of Euclid.

The truth of Euclidean geometry was seen then to have both empirical and philosophical grounding that could never be undercut. But that view suffered two shocks a century after Kant. The first and greatest was the discovery of non-Euclidean geometries at the end of the nineteenth century. Kant's own foundational account was widely rejected and the discovery gave rise to a plethora of mutually exclusive alternative answers to the foundational question. The second shock came in 1916 when modern physics was presented with general relativity, which adopted non-Euclidean geometry as the geometry of real, physical space-time. Several of the attempts to deal with this development built on earlier reactions to the advent of non-Euclidean geometries. One important way of dealing with the problem, later endorsed by Einstein, was to distinguish geometry as a formal enterprise from geometry as an empirical enterprise.[4]

In the years after the acceptance of relativistic physics, theorists came to see that physical science does not, even within the limits of fallibilism, provide a basis for any absolute knowledge of what exists (see Appendix, pp. 205–6); this is because what is known is dependent upon non-theoretical (e.g., practical) choices that the investigator must make logically prior to the stage of pure theory. Physicists have come to acknowledge that even the most widely accepted physical theory cannot guarantee (whatever the present state of empirical data collection) that no other theory will account for those data as well as the preferred theory. Physical theory includes an axiomatic system of geometry – but physical theory and observation cannot offer assurances that the system of geometry that physicists endorse is in any absolute sense the 'true' geometry of the world. There will be an inescapably *conventional* element to the system of geometry used in physics.

Poincaré (1905) was the first to argue that the system of geometry that physicists choose is inherently conventional. The simplest way to understand the point is to consider his parable of a two-dimensional world inhabited by a group of two-dimensional scientists. They do not know that we, from another world, are exerting a force on to their world that will shrink or expand two-dimensional rods as they are moved from one part of the two-dimensional surface to another (say expansion from south-east to north-west and contraction in the opposite direction). The scientists erroneously accept hypothesis (HG1) that *measurement of distance should make use of measuring rods, which remain rigid when moved* and they reject (HG1') that *measurement of distance should make use of measuring rods, which expand or contract when moved*. We see (HG2) that *the geometry of their world is Euclidean*. But their acceptance of hypothesis HG1 will lead them (e.g., by comparing radii and diameters of circles) to conclude (HG2') that *the geometry of their world is non-Euclidean*. If an unorthodox scientist in the two-dimensional world were to come up with the idea that there is a distorting force, and thus conclude that HG1 is false and that HG2' is true, there would be no way to prove it. The choice between accepting HG1 and HG1' is non-theoretical and logically precedes the task of fitting theory to observation.

A more thorough but slightly more complex way to see the conventional point is to imagine two different two-dimensional worlds. One world is a flat surface inhabited by group A, the other a hemispherical world inhabited by group B. We can picture a pair of intervals (line segments) in the flat world and a pair in the hemispherical world. We suppose that members of group A measure intervals in the normal way by assuming HG1 that measuring rods retain their length, and thus remain rigid, when moved. Thus members of group A conclude (e.g., by the method in the first example) HG2 that the geometry of their world is Euclidean. Members of group B assume HG1 and conclude, by the same method, HG2' that the geometry of their world is non-Euclidean. We next suppose that group A changes its method of measuring intervals to (HG1") that *measurement of distance should proceed by taking the projection of the intervals onto the flat surface* – which is clearly inconsistent with HG1, since HG1" would require the assumption that measuring rods expand as they move from world A to world B. Members of group A would then be able to conclude HG2' that the geometry of their world is non-Euclidean. It follows that one's choice between acceptance of HG1 and HG1" will determine whether one believes that one's world is described by geometrical system HG2 or HG2'.[5] In our own world this conventional element in the conclusions we draw about our physical space is irreducible and inextricably tied to the method of measurement.[6]

Poincaré's solution has also been interpreted as a way to salvage the Kantian view, because the 'free stipulation' of a system of geometry leaves the way open for that stipulation to be Euclidean.[7] This is particularly clear as Poincaré went on to argue that, even though the investigator has a choice between Euclidean and non-Euclidean geometry, the investigator will in fact always prefer the stipulation of Euclidean for its simplicity, which turns out to be false. However, modern physics can be constructed to conform to all known observations and still use Euclidean geometry (see Havas 1967: 140–1).

A decade before general relativity Duhem correctly argued against Poincaré that Euclidean geometry may not always be preferred. The Duhem thesis states that no subset of a theory-and-its-attendant-heuristic, such as an individual hypothesis, can be definitively tested (and rejected) as a discrete subset of the theory; negative results of an hypothesis test can show only that there is some flaw in the hypothesis, *or* in the laws of the theory *or* in the testing methodology, etc. Theories can only be properly viewed and treated holistically (Duhem 1954: 187). Once this is recognised, it becomes clear that the simplicity test must be applied to the geometric system-plus-physical theory whole. When this is the unit of appraisal, it might very well be that, despite the view that Euclidean geometry is simpler than any non-Euclidean system, insistence on the former can create so many additional complications in the attendant theory that the Euclidean-plus-physical theory whole may turn out to be less simple than the non-Euclidean-plus physical theory whole. This is indeed what happened. Poincaré's prediction of physicists' unwavering preference for Euclidean geometry was shown to be mistaken after 1916. But the error in no way reflects a flaw in conventionalism, since the mistake can be accounted for within conventionalism, e.g., as Duhem (1954, 1969) does.

Some conventionalists go much further and argue that, even after a physical theory is chosen, there is no univocal way to determine which axiom system of geometry should be chosen. Poincaré (1905) takes the argument in this direction, as do Hans Reichenbach (1958) and Adolf Grünbaum (1968). Grünbaum, for example, contrasts discrete and continuous manifolds and argues that modern physics conceptualises the four-dimensional manifold of space-time as continuous. In a discrete manifold the notion of a metric intrinsic to the manifold makes some sense: one can 'count points'. However, in the case of a continuous manifold, no such notion is applicable due to the equi-cardinality of all positive intervals of a continuous manifold (a result proved by Cantor). Thus space-time is what Grünbaum calls 'metrically amorphous'. While this view has been contested by Putnam (1975) and Friedman (1983, 1992), what has not been contested is that one cannot make a theory choice in physics without a prior practical decision about methods of measurement.[8]

In sum, it is a misapprehension to believe that physical theory is inherently capable of producing knowledge not dependent on non-theoretical decisions, especially the method of measuring length. This misapprehension has led to unfounded expectations of social science by naturalists. Hence many assume that physical theory produces unambiguous theory choice and a single set of ontological commitments to which physicists adhere. They infer, *qua* naturalists, that social theory should be able to do the same. But, as noted, even within a single area of theoretical enquiry, like relativistic space-time theory, which lies at the heart of modern physics, the data do not force upon the investigator any single theory or ontology. Alternatives are consistent with the discipline's best theoretical knowledge. Accepted physical theory is not absolute, even beyond the widely accepted qualifications of fallibilism. Rather, there is an inherent element of conventional choice that the investigator must make. This conclusion casts no doubt on naturalism, since the need for a conventional choice remains whether one is discussing IR or physical theory.

Incommensurability, radical underdetermination and hermeneutics

Among the most common explanations for lack of progress in the social sciences among naturalists are IP and RU, and among anti-naturalists it is the role of interpretation in HT. While IP and RU have been proffered as accounts for dissensus in IR, none succeeds in answering all three questions stated above on the first page of this chapter. In particular, neither IP, RU nor HT answers question (ii), how there has been consensus in physical theory, nor do they answer question (iii), how there has been approach-toward-consensus in within IR, e.g., in the area of DP theory? The HT has trouble accounting for the differential successes in terms of cumulation and approach-to-consensus of the various social sciences.

Incommensurability

The IP thesis is generally believed to undermine traditional accounts of science that see theory choice as objectively – or at least as intersubjectively – valid. The outlines of Kuhn's philosophy of science, which includes the IP thesis, are well known. Kuhn (1962) rejects the logical positivist and logical empiricist accounts of science, which dominated the middle of the twentieth century, by arguing that they were entirely disconnected from the actual history of science. Theory choice is not a regular feature of scientific activity. Science proceeds quite dogmatically within a dominant 'paradigm' (later termed 'disciplinary matrix'), which includes methodological rules and procedures for testing hypotheses. The dominant theoretical 'exemplars' have various anomalies, which are not ordinarily taken as falsifying the dominant theory. There are moments of crisis where the activities of 'normal science' give way to genuine theory choice. However, the choice is not made on the basis of philosophical, rational and logical grounds of the sort specified by logical empiricists. Rather, according to Kuhn, the outcomes of the choices are driven by sociological factors.

There are two reasons most often cited with regard to why paradigm choice is not rational. One is that, following IP, the paradigm itself contains the methodological rules, normative prescriptions and testing procedures. There is no set of rules external to the paradigms that permits comparative appraisal. That is, there are no rules or standards *outside* the competing paradigms which could be taken as 'objective' and employed as a basis for unbiased choice between them.

A second reason, and perhaps the most oft-cited, why theory or paradigm choice is said not to be based on rational grounds is that the logical empiricist methods of science and comparative theory appraisal require a theory-neutral language in which to express basic observations. But, the argument goes, observation terms, which may appear the same in both theories (e.g., 'mass' in Newtonian and relativisitic physics) do not have the same meaning in the different theories. Their meanings are given by the theories in which they are embedded. Thus there is no true theory-neutral language and theories are incommensurable. There is a 'gestalt switch' as scientists move from one paradigm to another. Those who advance this position often use analogy of the 'duck–rabbit' invoked by Wittgenstein (1953) and Hanson (1958). One may see the diagram as a duck or as a rabbit, and one may switch back and forth between the two but one may not see it as both at the same time. This argument, discussed on pp. 66–76, was developed in the philosophy of science by Hanson (1958), Kuhn (1962) and Feyerabend (1978). The IP thesis implies that when there is lack of consensus in a particular field it may be a result of the impossibility of evaluating contending theories against one another. (See Hunt's 1994 useful account.)

With respect to the first argument for IP, Kuhn is notoriously vague about what constitutes a 'paradigm' in *The Structure of Scientific Revolutions*.[9] In his subsequent writings Kuhn has moved away from use of the term 'paradigm' altogether, distinguishing it into more than one concept, including 'disciplinary matrix' and 'exemplar'. 'Paradigm', however defined, is a broader term than 'theory', since the former includes the latter (but not vice versa) along with various methodological principles. But presumably it does not include everything, such as

i) *principium contradictionis*;
ii) other elements of deduction;
iii) elementary number theory;
iv) the least philosophically problematic areas of statistics and probability (at least when certain conditions of data-collection are met); and
v) other methods or tools of analysis.

The question amounts to this: what else is included in the 'other methods' category? If enough 'other methods' are included, then theory or paradigm choice may well be possible, at least much of the time. While there are many ways to attack the IP thesis, much of the problem IP poses is that it claims to prevent rational, scientific comparison. The degree to which this is a genuine problem is going to turn on how the unit of comparison (e.g., 'paradigm') is defined.

It should be noted that IP, even if it is held to account for the lack of consensus and progress in IR and various other social sciences, does not answer question (ii) above, how the lack of progress in IR can be understood while there is much greater progress in the physical sciences. IP sees both sorts of disciplines as lacking any external rules or standards, or a common, theory-neutral language for judging theories and communicating across paradigms. Much effort has been expended by Kuhn, Feyerabend and their supporters to show how IP does account for the history of natural science, particularly physics. Those accounts have been vigorously disputed on historical grounds. IP denies the presupposition of question (ii), namely, that there has been rational, cumulative progress in physics. Paradigm choice, or disciplinary matrix choice, is a result of non-rational sociological factors. Space limitations preclude the sort of detailed enquiry into the history of physical theory that would allow adjudication of the competing sociological and rational accounts. This study recognises the extensive difficulties that have been cited in the literature that would allow one to conclude that there is no more cumulation and progress in physics than there has been in any of the social sciences.

With regard to the second argument, the discussion in Chapter 3 shows the difficulty of supporting the theory-ladenness of observation thesis, that there is no theory-neutral language capable of permitting an adjudication between or comparative appraisal of two competing theories (or 'series of theories', in Lakatos's terms, or 'paradigm', in Kuhn's terms). Moreover, Fodor (1984) and Greenwood (1990) have shown quite persuasively that the charge of theory-ladenness is based on a conflation of several concepts which, once recognised as distinct, prevent further confusion and obviate the need to withdraw from 'objectivity' or 'intersubjectively valid' conclusions about theory choice.[10]

Greenwood's argument is significant in this context because, even though he accepts a limited form of the principle of the theory-ladenness of observation, he shows successfully that even if that principle is true, IP does not follow from it. Greenwood identifies two interpretations of the theory-ladenness principle (which he calls 'theory-informity1' and 'theory-informity2'). The first version (1990: 561–2) holds 'that observations can always be interpreted according to explanatory theories which are the object of observational evaluation'. But he observes that this formation does not allow inference to IP, rather, it requires use of IP. It 'presupposes that observations can never be made that enable us to objectively discriminate between competing theories. For only if this is the case is it true that observations can always be interpreted according to different competing theories' (1990: 562).

The second sense of the theory-ladenness principle is drawn from claims that 'we often need independently supported *exploratory theories* in order to make observations that are critical in the evaluation of explanatory theories' (Greenwood 1990: 562). These exploratory theories allow investigators to 'see' how the purported behaviour of electrons or DNA molecules appear to them on experimental apparatuses. Greenwood argues that the use of an exploratory theory makes possible 'the observational *mensurability* and *commensurability* of competing explanatory theories: it enables scientists to make critical observations in favour of one theory over its rivals. Incommensurability is a product of the absence (or temporary poverty) of exploratory theories at certain historical periods' (Greenwood 1990: 562), as when theories of the structure of proteins were 'observationally incommensurable' as a result of the absence of an apparatus (X-ray diffraction) that would allow observational evidence for or against any of the competing theories.

There is an extensive debate on the acceptability of central tenets of Kuhn's account, particularly IP. The above remarks are intended to show specific flaws in IP as an account of the physical sciences. With respect to trends in the debate over IP, Kuhn and other defenders have had to weaken IP significantly in order to defend it from well-founded criticisms. These attempts have diluted IP to the point that recent defenders, while retaining a sense of 'incommensurability', indeed allow for theory or paradigm 'comparability', as is clear in Sankey's (1994) attempt to salvage Kuhn's thesis. Thus the familiar form of IP encounters severe problems answering question (ii). As noted above, even if physicists do not fully agree all the time, the history of approach-to-consensus when new physical theories are introduced is vastly different from the history of continuing dissensus in most areas of IR. Though there are exceptions, such as DP studies (discussed below on pp. 189–92), the failures of consensus endure even in many areas of IR where most of the authors share the naturalist's enthusiasm for the methodological analogy of the physical sciences.

Radical underdetermination of theory by data

The doctrine of underdetermination of theory by data, in its radical or other forms, is advanced by many social scientists and philosophers (Kyburg 1990a). It is generally held to imply various forms of relativism, such as that any body of evidence always offers equivalent support for more than a single theory. In its extreme forms the charge of relativism is most plausible. A classic statement of the RU thesis by Quine, in 'Two Dogmas of Empiricism', where he says that the investigator may hold on to any favourite statement or theory 'come what may' as new evidence accrues (Quine 1953: 43). That is, one may always concede other beliefs and reconcile one's favourite theory with the accepted evidence. Quine's view leads to a form of relativism, since all theories can be reconciled with the evidence. Even more relativist is Quine's view that the given theory and its rivals are not only both supported by the evidence, but are *equally well* supported by it. This view has been endorsed by a variety of philosophers of science, including Feyerabend.

Quine's position has come under sharp attack. Laudan (1990) argues persuasively that relativists' support of these versions of RU is based on equivocation on the different versions of the underdetermination thesis (he distinguishes half a dozen). Duhem's original and much more modest formulation of the underdetermination thesis states that there is always a rival (or perhaps an infinite number of rivals) to a given theory unrefuted by the evidence supporting the given theory. But even if there is an infinite number of unrefuted rivals, there may also be an infinite number of theories refuted by the evidence. (Duhem's notion of what is falsified is a 'theory' in the broad sense, something like Kuhn's 'disciplinary matrix'.) Duhem's more limited formulation of the underdetermination thesis is cogent and is fully capable of supporting Popper-style *modus tollens* falsification. It does not follow from it that any theory whatever may be retained in the face of any set of observations whatever. Moreover, the given theory may even be preferred over the many unrefuted rivals on rational, though not strictly deductive, grounds. Only the stronger, more suspect, formulations of underdetermination, such as Quine's RU, lead to relativism or to the claim that scientific theory choice must ultimately be understood sociologically rather than rationally.

Quine's RU holds that any body of beliefs can be salvaged in the light of any newly accepted observation or statement if one makes the right adjustments to the body of beliefs (and in the case of scientific theories, if certain auxiliary hypotheses are added). The adjustments are not, however, all equally plausible. Other philosophers of science have accepted the purely deductive consequence of Quine's claim and sought other grounds for rejecting statements, theories or paradigms in science when a new observation is inconsistent with the previous theory or body of beliefs. Lakatos's attempt to distinguish the sorts of changes in old theories into 'progressive' and 'degenerating' is one of the most well-known examples.

Goldman (1999: 243–4) points out that RU makes use of the possibility that two theories may have an equal number of confirming observations. He points out that at least for Bayesians, one might reject the Quinean claim that two theories with equal confirming instances are thereby equally well confirmed. For Bayesians they may be dissimilarly confirmed because the two theories may have had distinct prior probabilities.

While the doctrines of IP and RU might in principle help to explain why there has been little progress in the social sciences, both have, as just noted, come under serious attack. But RU, like IP, accounts for failure of progress in IR and social sciences in a way that *prevents* it from answering how the physical sciences have progressed, since the underdetermination features of social science that impede progress are also present in the physical sciences. RU cannot answer questions (ii) and (iii), where there have been examples of approach-to-consensus.

The hermeneutic tradition

The HT, as described in the previous chapter, is fundamentally opposed to the naturalism and the analogy on which it is based, that between the natural sciences and the social sciences. The social sciences are seen as explanatory but not predictive. And they are explanatory in ways very different from the natural sciences. Moreover, the social sciences are rule-governed in a double sense, since the enquiry by theorists into the actions of people, groups and states follow rules of enquiry and inference, while the behaviour of the agents under investigation is governed by a different set of rules. Investigators seek to interpret the actions of the agents. However, according to the HT, each theoretical interpretation is relativised to selection, perspective, a set of unspecified assumptions, etc. (Bohman 1993: 110). There is a circularity involved in every interpretive effort. Hence, there will not be agreement of two different investigators at different places and times, since they will make use of different contexts of reference for the objects in the interpretation. The result is that social science cumulation and approach-to-consensus are not possible. Furthermore, they make it clearer than defenders of IP and RU that the social sciences and the physical sciences are entirely dissimilar in their structure of their theories. They do best at explaining the different histories of approach-to-consensus in the physical sciences versus the permanent state of dissensus in the social sciences.

The problem with the HT's attempt to answer the three questions posed at the outset of this chapter are concentrated on the last, as the HT cannot account for approach-to-consensus, where it exists, in the social sciences. Neo-classical economics has produced much more approach-to-consensus than could be explained on an HT account, where all theory choice turns out to be choosing between competing interpretations, and all such choices are subjective. In the case of IR, furthermore, there does seem to be emerging agreement in some areas of enquiry, most notably in DP studies. The HT would seem unable to account for such agreement. As noted in Chapter 5, Bohman's HT-oriented interpretive account of the social sciences holds that the new approach is to take the history and practices of the social sciences as central to any methodological and meta-theoretical analysis. However, his examples are ethnomethodology, theory of communicative action and rational-choice theory. While the latter includes economic theory, he does not engage actual theories that have proven successful. He looks at general aspects of rational choice reasoning and concludes that it cannot be predictive. But he does this without looking at the ability of the theories to predict, even though he says that the new philosophy of social science and the actual practices of social scientists should play a central role and serve as key evidence in theorising about the social sciences. The difficulties of the HT in IR are seen most clearly only by an examination of the development of DP studies, to which attention is turned presently (pp. 189–92 below). Ultimately the relativism of IP, RU and HT are insupportable because of their inability to explain the empirical record of theorising in areas with different histories of success or failure in achieving approach-to-consensus. This accords with Donald Davidson's view that such relativism is incorrect, because massive error is unintelligible. As Davidson says, '[i]n order to communicate, most of our beliefs must be true', and adds, 'conceptual relativism presupposes an incoherent "final dogma of empiricism," that uninterpreted experience is uncontaminated by interpretation' (Davidson: 1984: 199).

Natural experience, social experience and the conventional account of consensus and dissensus

An explanation of both the history of agreement in the physical sciences and the lack of agreement in social sciences like IR is possible by means of CC when conjoined to further considerations regarding 'epistemic access' to the objects of physical and social enquiry. Although this book defends a fairly strong version of naturalism, the social sciences are of course not identical to the natural sciences and thus there must be points at which the analogy breaks down. One point, noted above is of course the divergent histories of the natural sciences and social sciences (and how this affects motivations for them). A second break-down point is that of the likelihood or degree of difficulty of reaching discipline-wide conventions of methods and measurement.[11] The second difference helps to explain the first. This section outlines an explanation of why the second difference arises, in terms of the relationship of scientific theory to pre-theoretical experience.

The key difference in the difficulty of reaching consensus in the natural sciences and social sciences is that, while people believe themselves to have direct perceptual awareness of sticks and stones, no such claim is made for regimes, states, power balances or perhaps even foreign ministers. Many philosophical accounts of perception capture precisely this. To take one example, Roderick Chisholm has stated his basic epistemological principle in an evolving series of formulations over the past four decades. The most recent version is as follows:

> 1) For every x, if x senses an appearance that is red, then it is evident to x that he or she senses an appearance that is red; and 2) Being appeared to in a way that is red tends to make it probable (confirm) that one is sensing an appearance of an external physical thing that is red.
>
> (Chisholm 1997: 39)[12]

This clearly links the most basic perceptual experiences to beliefs in external *physical* objects. Furthermore, Chisholm's principles are restricted to perception of physical objects and do not offer any parallel treatment of human nature, states, regimes, power balances or other social entia.

While it is possible to perceive social objects, as argued in Chapter 3, one's perceptual access to them is fundamentally different from one's access to physical objects. The purpose of citing Chisholm here is to show the difference between perceiving physical and social objects. The many criticisms over the past forty years of Chisholm's account (e.g., by Alston 1990, Dretske 1979, Lehrer 1986 and many others) criticise many features of his account but they do not include attacks on his differential treatment of physical and social objects. Thus the extensive discussion over the years of Chisholm's well-known account is one indicator of the wide philosophical agreement on the different sorts of access that an individual has to physical versus social objects.

In the social sciences there is disagreement not only over the basis of social concept formation, but also over the most appropriate methodological assumptions. While there is some debate in the natural sciences, the breadth and depth of disagreement is not parallel to that in the social sciences. To take physics as the clearest contrast, as shown above, physicists widely accept the assumption that some physical bodies, including measuring rods, remain rigid when moved. It seems likely that one of the reasons that a consensus has been possible among physicists is that they are dealing with physical bodies and, as philosophical accounts reveal, people have fundamental beliefs about those bodies arising from their most basic (or least theory-laden) perceptions.

The situation that confronts IR theorists and many other social scientists is quite different. Political realists in IR use one set of methods and measures for their theories while liberal institutionalists use a different set. Thus far there have not been any basic methodological assumptions, including measure-stipulations, in many areas of the social sciences which are parallel in the way they accord with pre-theoretical experience and which have garnered consensus. The CS thesis, however, allows for the possibility of such consensus in some social sciences, as there has been in the area of DP research. The philosophical literature dealing with accounts like Chisholm's suggests that one of the reasons that disagreement persists over the acceptance of a single metric (like 'state power') in IR, while physical scientists have been much more able to agree on measure-conventions, arises from the difference in the way concepts of physical versus social entities relate to people's pre-theoretical experience. The proponent of CC would expect differentials in patterns of approach-to-consensus in the physical sciences and social sciences because of the different sorts of roles played by perception in the case of physical and the social sciences.

The measure-stipulation and 'power'

Classical realists see human nature or the nature of the state as the driving force behind conflict behaviour in world politics, while neorealists emphasise the anarchical structural order. For Waltz, any political theory must understand what a political structure is. He regards the defining characteristics of a political structure to be the ordering principle (anarchy or hierarchy), the functions of the units constituting the structure, and distribution of capabilities among the units (one, two, three or many poles of power). For either classical or structural realists, frequency, intensity and persistence of inter-state war are often used as measures of conflict. But the most consistent unit of measurement, whether denominated 'distribution of capabilities' or something else, is 'power'.

Idealists, functionalists and liberal institutionalists look at the same historical record but see a different picture. For example, while neorealists see the development of the Bretton Woods system as a result of the preponderance of American power and the ability to coerce that accompanies hegemony, opponents of neorealism see regimes and institutions that allowed the US to choose to pay large costs to create the stable monetary order. Many opponents of neorealism have included normative imperatives in international politics and reject any view that permits the ends to justify the means. They likewise reject the state as the permanent unit of action in world politics. Rousseau, Kant, Bentham and Wilson saw current institutions of their eras as impediments to peace and regarded the creation of a new set of institutions more suited to promoting lasting peace as a genuine possibility.

Recent institutionalists emphasise the centrality of the notion of 'regimes' in world politics and the way in which systems with one set of regimes will exhibit different behaviour (e.g., in patterns of cooperation) than systems with significantly different regimes. Because regimes permit states to adjust policy with less fear of having to incur high costs (e.g., by rivals taking advantage of their generosity) states are more likely to adjust policies accordingly and more cooperation results.

Contemporary institutionalists see a variety of circumstances under which China, France and Uruguay will choose to maximise gain without being limited by concerns that rivals will simultaneously gain. Thus states may seek to maximise absolute gain in many instances when political realists expect them to maximise relative gain. Institutionalists like Keohane tend to measure cooperative behaviour by examining what states do in the areas of trade and finance rather than frequency of war. Their units of measurement are often technological status, efficiency of production or comparative advantage.

Political realism and liberal institutionalism have always differed sharply over the possibility of future cooperation in IR. For political realists, if humans are fundamentally at odds with one another, the best that can be hoped for is to limit the extent or frequency of violence. There remains serious doubt, at least outside of a hegemonic international system, as to who could possibly do the limiting. For liberal institutionalists, improvement of human society is possible. Can the measuring rod of power remain rigid when moved from security to trade affairs? Although his central concepts are all drawn from political economy examples in telecommunications, oil, textiles, etc., Keohane (1984) says that his theory can be expanded to account for international security relations.[13]

A similar point about divergence can be made when one compares other research traditions in IR, such as Marxism, bureaucratic/domestic politics, interdependence, reflectivism, etc. The inability of IR theorists to agree is evident when any two or more competing research traditions are compared. The effects of the measure-stipulation can help explain that pattern. The discussion here of realism and liberalism is just one instance of the dissensus in IR.

The democratic peace, conventions and the 'measure-stipulation'

Though charges that the social sciences have not exhibited progress and cumulation are often well founded, there are areas where they do not apply. The CC account of IR theory argues that there is the possibility of cumulation and approach-to-consensus in a way not unlike that in the natural sciences. This section and the next two sections consider one example of progress, studies of the relationship between 'democracy' and 'peace', which many scholars regard as the most promising in this regard. This example of progress is highly significant, since it deals with a core question of IR. As Harvey Starr puts it, 'Given that war has been perhaps the single most central concern to students of international relations across history – and certainly to Realists – uncovering one factor, variable, or set of conditions that is associated with the complete (or almost complete) absence of war is a stunning achievement' (Starr 1997: 114).

Kant of course posed the theoretical hypothesis that a system of republican states would lead to a more peaceful world because citizens would play a role in decisions about war and peace. They would engage in free trade, which would create interdependencies that would make war economically unwise. And they would form a league or federation to promote peaceful norms and methods of conflict resolution. A 1972 article by Dean Babst made the empirical claim that 'no wars have been fought between nations with elective governments between 1789 and 1941' (Babst 1972: 55; see also Babst 1964). R.J. Rummel 'was most responsible to calling attention to the phenomenon of the democratic peace' (Chan 1997: 61), as he first subjected the claim to rigorous testing in the political-science literature (Rummel 1979, 1981, 1983, 1992, 1997). Works by Small and Singer (1976) and Michael Doyle (1983a, 1983b, 1986) further established several DP hypotheses as central to the contemporary study of IR.

The current wave of DP studies began with authors disagreeing about whether it is true that democratic states are more peaceful than non-democratic states. Some prominent mainstream IR scholars questioned the empirical claim (e.g., Small and Singer 1976). As debate proceeded, there were various clarifications and, in some cases, conventional decisions were adopted by the majority of scholars in the field. This allowed a large degree of consensus to emerge over twenty-five years on several claims: 'that regime type has no effect on conflict behaviour' (the null hypothesis) is false; that 'democracies are more pacific than non-democracies' (the monadic hypothesis) is false; and that 'democracies never, or almost never, go to war against one another' (the dyadic hypothesis) is true.

While there is consensus on some behavioural hypotheses, there is still much disagreement about the causal mechanisms that produce the observed effects.

Recent studies have extended the scope of DP studies to include other elements of Kant's theory, especially the role of liberal trade policies and international organisations. Additional DP hypotheses and corollaries have been formulated and tested, though with more mixed results. Examples include that democracies do not initiate as many wars as non-democracies; that democracies are less likely to intervene covertly in the affairs of other states, particularly other democracies; that democracies are less prone to fight extra-systemic or colonial wars; and that wars fought by democracies are more limited and have fewer casualties than those of non-democracies.[14]

Both proponents and opponents of DP hypotheses have come to accept certain methodological claims, e.g., that the pacificity of democracies, liberal states, autocracies, great powers, etc., should be measured by the proportion of wars they fight. The convergences on the null, monadic and dyadic hypotheses just noted are not the sorts that supporters of the HT, RU and IP would expect. There has indeed been dialogue between proponents of DP hypotheses and empirical critics, as well as a group of critics in the constructivist or postmodern camp. This can be seen by examining the discussions of the terms that scholars have used to formulate the key hypotheses and the approach-to-consensus on their definitions.

In terms of the dependent variable, scholars have debated what should be counted: mere absence of war, stable peace, interdependence, etc.? Should one count threats of the use of force that do not subsequently escalate to violence or covert actions that are shielded from the public? Studies have largely adopted the convention of 'absence of war' as the measure for core DP hypotheses, although lower levels of violence have been studied for several reasons, e.g., to gain deeper insight into the mechanisms of democracy that produce peace and which specific features of democracy lead to peace, or to use some statistical methods because wars are not frequent enough for meaningful testing. Other questions about the dependent variable have been raised, such as whether peaceful democratic norms ought to preclude also domestic (civil) war, whether conflict should be measured qualitatively or quantitatively and, if the latter, how much violence is permissible before one counts the event as a war?[15] The choice of which particular notion to adopt as the dependent variable for what are regarded as the core hypotheses is largely a conventional one. Similarly, the quantitative question of how much violence is required for an incident to be regarded as 'war' has been answered, following Singer and Small (1972), by a conventional decision to place the threshold at 1000 battle fatalities.

In terms of the independent variable, one must ask what sorts of regimes are more pacific. Some states are liberal but not democratic (Britain before 1832), while others are democratic but not liberal (the Confederate States of America). So distinctions must be drawn between states that have constraints on their leaders, have accountability of their leaders, are liberal (defined in terms of political liberties and rights), are republican (equal rights among the – possibly minority – group of decision-making citizens), are democratic (defined in terms of voting rights), and are libertarian (defined by Rummel (1983) in terms of a combination of political and economic freedoms). If 'democracy' is chosen, as it often is, then, in formulating and testing DP hypotheses it is important to clarify what features of a state must be present for the state to be included as a democracy, e.g., mass participation, protection of the rights of minorities, individual liberties, etc. Once the distinctions are drawn and the preferred term selected then, as with the dependent variable, there is still a question of whether it is to be investigated by means of a qualitative measure (democracy versus non-democracy) or quantitative measure (degrees of democracy/autocracy). The two approaches yield different results. Russett (1993), for example, points out that the less liberal democracies in ancient Greece were more war-prone. There is also a difference between initialising democracy (building institutions) versus consolidating it (creation of a civic culture) as some hold that only the latter is relevant for the DP hypotheses.

Overall, the past thirty years of DP studies has led to a large degree of consensus. There are many examples of authors who began with different conclusions, stemming from divergent ways of operationalising their terms, altered their definitions and methods, and consequently came closer to agreement. As Chan (1997: 62) puts it, 'All this research contributes to a vigorous debate – surely, one of very few areas in international relations scholarship in which, notwithstanding substantial disagreement, considerable progress and excitement have resulted from an iterative process of criticism and response'. He says also, 'the democratic peace proposition is one of the most robust generalisations that has been produced to date by the [quantitative] research tradition' (Chan 1997: 60). Gleditsch says, 'the importance of democracy lies in it being a near-perfect sufficient condition for peace' (1995: 297). Russett says that it is 'one of the strongest nontrivial or nontautological generalizations that can be made about international relations' (1990: 123). Levy says that the dyadic hypothesis is 'as close as anything we have to an empirical law in international relations' (1989: 270). Beck, Katz and Tucker (1998: 1274) say that 'that democracies do not go to war against one another' is 'one of the most celebrated propositions in the IR/IPE literature' and that it 'has been confirmed by myriad empirical studies' (1998: 1274). Starr (1997) concurs that a consensus has been reached on some of the central questions that DP studies have investigated.

Starr (1997) says that Richardson (1960), Rummel (1968) and Small and Singer (1976) denied the monadic hypothesis. However, Rummel later published a paper supporting it (1983), other authors took up the question and Chan (1984), Weede (1984), Garnham (1986) and Maoz and Abdolali (1989) 'provided evidence against Rummel's conclusion and reinforced a consensus that democracies have generally not engaged in less war than' non-democracies (Starr 1997: 112–13 n. 1). Starr notes that there is some dissent (Ray 1995: ch. 1). With regard to the dyadic hypothesis 'that democracies do not (or only rarely) go to war against one another', Starr (1997: 112–13) says, '[t]he empirical findings across a number of studies have produced a consensus in support of this hypothesis'. Even critics of DP hypotheses also see a general consensus on at least some DP hypotheses. Realist opponents like Farber and Gowa (1997: 394) acknowledge a 'strong consensus on two related issues', namely, the dyadic hypotheses about war and about other lower-level disputes.

Advocates of the HT and RU would not expect this sort of convergence. However, with respect to IP, critics of 'progress' like Kuhn could account for the growing strength of the proponents' side in the debate and still deny genuine 'progress' by explaining it by means of sociological and historical rather than rational influences. Such critics could still maintain that the terms used by the proponents and opponents have divergent meanings – the paradigms are incommensurable – and thus the two sides do not truly engage in 'dialogue'. Thus even if one side gains numerical strength, the two sides are still arguing past one another. So a more telling evaluation of progress would be to take the key strands of the DP debate and see how the two sides interact with one another – how critics attack DP hypotheses and how defenders respond to the attacks.

Empirical critique of democratic peace studies

Definitions

The dialogue on the definitions of key terms in the DP literature is one example of the progressive and cumulative nature of the field. Scholars have come to agree on certain conclusions related to conventions on various independent and dependent variable measure-stipulations. As studies have reacted to one another, authors have indeed adopted some of the previously published definitions, which has resulted in greater agreement on these hypotheses.

There are numerous examples of genuinely progressive interaction between scholars over the question of the definitions of key terms. Russett's work provides significant examples of this process. Russett, in his paper with Zeev Maoz, develops an institutional measure of democracy: *Regime Type* = *(Democratic Score–Autocracy Score)* × *Executive Power Score* (Maoz and Russett 1993, see also Tures 2001.) This measure was used for several years, until Thompson and Tucker (1997: 448) pointed out that those states with greater concentrations of executive power are more likely to be coded as 'democracies', even though such states are commonly regarded as less democratic. In the wake of this comment, in subsequent work Russett dispensed with that measure (Oneal and Russett 1997: 274). This is the type of interaction that characterises the DP studies as genuinely cumulative and progressive.[16]

Progress arises from 'an iterative process of criticism and response' (Chan (1997: 62). And the iterations continue. Russett, in a paper with Oneal, once again acknowledges room for improvement even in the revised measure of 'democracy' based on published replies (Oneal and Russett 1999b). He considers criticisms of his revised definition (*democracy score* – *autocracy score*) presented by Kristian Gleditsch and Michael Ward (1997), according to which the measure has 'the virtue of being symmetric and transitive' but also allows that 'the relative importance of its components is unstable over time' (Oneal and Russett 1999b: 12 n. 30). Oneal and Russett acknowledge that the shortcomings of their definition and the need to improve it in future studies.

Theoretical critique: interests versus joint democracy

Farber and Gowa

Even critics of the DP hypothesis engage the literature, use similar definitions and methods, or at least start with them, and build their critiques by arguing for the need for adjustments in them. The realist critique by Farber and Gowa (1997) is an excellent example. They argue that 'interests' account for the pattern of dyadic war and peace better than 'regime type'. However, their argument does not proceed as the Kuhnian IP account would expect by offering a distinct set of terms (or a set of terms that take on different meanings because they are embedded in a distinct theory). Rather, as they are used by many critics, the terms 'democracy', 'autocracy', 'anocracy', 'militarised inter-state dispute', and 'war' have the same meanings as they do when used by supporters of DP hypotheses, since both sides of the debate start with the same data sets, especially correlates of war (Small and Singer 1976) and polity II and III (Jaggers and Gurr 1995). Even in arguing for their preferred realist variable, i.e., 'interests', Farber and Gowa make use of the same data set that DP proponents use for many of their tests, especially Singer and Small's correlates-of-war data. When deviating from the definitions used in the data sets, authors accept the burden of defending their deviations. (When Layne (1994) criticised the definition of 'democracy' in DP studies without offering an explicit definition of his own, the failure did not go unnoticed; see Russett 1995: 167.)

Even though most proponents are liberals and many critics are realists, the two groups do not in general misunderstand one another. The most cogent works by both groups can be seen to use the same terms and to advance the debate to greater and greater degrees of precision and depth. Farber and Gowa are careful to make sure that their 'data yield the results consistent with those of previous studies' (1997: 404). They are clearly building upon the work of DP proponents as they present their critique, in which they argue that 'security interests' account for patterns of war and peace better than 'democracy'.

What accounts for Farber and Gowa's different results is that, while using the same 1820–1980 data that Russett and other DP proponents use, they disaggregate it by dividing it into three periods: prior to the First World War, the inter-war period, and after the Second World War. They show that the DP hypotheses do well in the post-1945 period but poorly in earlier periods, especially pre-1914. There were so many new states in the system after the Second World War that the results derived by others based on aggregated data are unduly biased by the particular systemic conditions after 1945. For example, in the dyadic analysis of the 165 years from 1816 to 1980 there are 284,602 dyads (for which there are complete data). However, the thirty-five years following the

Second World War constitute just over a fifth (21.2 per cent) of the number of total years but includes nearly two-thirds (65.7 per cent) of the dyads. Thus when the results are aggregated, any factors stemming from the post-1945 system will disproportionately affect the statistical tests. The point here is not that the DP hypotheses are true or false. It is rather to show that proponents and opponents are not just talking past one another; they are speaking the same language and building upon one another's methods and developing new results to which the other side will be expected to respond.

The responses to the serious challenge of Farber and Gowa offered in Oneal and Russett (1999a, 2001) and in Russett and Oneal (2001) primarily deal with causal and theoretical issues. Farber and Gowa use alliances as their indicator of 'common interests' when testing their argument that 'common interests' explain conflict patterns much better than does 'democracy'. However, Russett and Oneal point out that alliances are not independent of democracy; ideology is often a reason for states to choose one another as alliance partners. This is not to say that geostrategic reasons have no bearing on conflict behaviour. Oneal and Russett concentrate on democratic and non-democratic regimes rather than anti-communism. (But the case would certainly be even stronger if one included communist/anti-communist regime ideology, which had a significant effect on Cold War alliance formation.)

Gartzke

Erik Gartzke (1998) also argues that common interests are more responsible for peace among democracies than their democratic nature. In his view, democracies have an affinity for one another that leads them to have overlapping rather than conflicting interests. Hence, the supposed effects of democracy in constraining the escalation of conflicting interests are greatly overstated, since there are no such conflicting interests and so not much to constrain. Gartzke argues that all of the proposed causal mechanisms that attempt to explain the low incidence of conflict among democracies mistakenly assume that there have been potential conflicts between democracies. Hence, there has not been as much 'prevention' of conflict or escalation, since there has been little to prevent.[17] Democracies' rarely conflicting preferences are what lead to lack of war, not anything about their democratic nature.

Gartkze uses UN General Assembly voting as a measure of 'affinity' or 'similar preferences', which some regard as superior to Farber and Gowa's 'alliance' measure, since there are many cases of states working together without formal alliances. It seems obvious that if democracies have greater affinities with one another, and states with greater affinities and similar preferences do not fight, then it would seem that democracy is still playing a crucial role. But what Gartzke argues is that the factors that 'cause' democracy, like 'ecological, material, or cultural factors' are the causes of the shared preferences among democracies (1998: 11). Russett and Oneal deal with the theoretical question of the causal role, arguing that democracy has causal effects on UN voting and alliances. The important observation here is that Russett and Oneal take up the challenge of 'preference similarity' and, using Gartzke's preferred UN General Assembly voting as the indicator, they respond to the criticism by testing the DP hypotheses using 'common interests' as one of the control variables.

Methodological critiques: dirty pool

In 1993 Russett published an influential book, *Grasping the Democratic Peace*, and a paper with Zeev Maoz (Maoz and Russett 1993), both of which focused on the pacific effects of democracy. Spiro (1994) criticised them on methodological grounds, especially the analysis of time-series data. Russett (1995) responds directly to Spiro. The papers by Oneal and Russett (1997) revise and extend the analysis (e.g., by bringing in 'trade' as an explanatory variable). The time-series techniques are criticised by Beck, Katz and Tucker (1998) and later by Green, Kim and Yoon (2001) on different grounds. Russett and Oneal respond directly to the critics in subsequent publications (Oneal and Russett 1999a, 2001 and Russett and Oneal 2001) and make adjustments to their analyses that do not render their programme degenerative. The interaction of DP defenders and critics provides a powerful example of the progressive nature of the debate.

In the first case, Spiro (1994) criticises the DP defenders' use of time-series analysis. Spiro argues that Russett's (and Maoz and Russett's) use of the pooled time-series method is illegitimate because the cases are not entirely independent; peace between the US and Canada in 1994 is not causally independent of the peace between them in 1993. Russett (1995) responds to Spiro's attack by offering an alternative method. He uses cases that are not dyad-years but are dyadic relationships over the entire period during which a relationship remains stable, which Russett terms 'dyad-regimes'. The entire two-century period of peace between the US and Canada becomes a single case. This avoids any charge of inflating the number of peaceful dyad-years with large numbers of observations that are not independent. Russett presents an alternative analysis of 'war' for the post-Second World War period, as well as analyses where the dependent variable is 'use of force' and 'disputes of any sort'. In the case of 'war', there were zero wars between democracies and 169 regime-dyads where there was no war. Of the 1045 regime-dyads that did not have a democracy, there were thirty-seven wars (3.4 per cent) among them. Russett regards the objections raised by Spiro as constructive, stating that the critique 'forced me to devise new tests. The result is that the evidence for the democratic peace is stronger and more robust than ever' (Russett 1995: 175).

The second example of criticism is that of Beck et al. (1998). In their 1997 paper, Oneal and Russett use logistic regression of pooled time series of state-year dyads in arguing that 'democracy' and 'interdependence' are significant factors in reducing international conflict (rather than just 'democracy' as in Russett 1993 and Maoz and Russett 1993). Beck et al. charged that Oneal and Russett encounter problems because time-series cross-sectional data, especially those using binary variables, undercut the assumption that the observations are independent. 'Since it is unlikely that units are statistically unrelated over time ...' binary time-series cross-section observations 'are likely to be temporally dependent' (Beck et al. 1998: 1260–1). These critics hold that corrections should be made for temporal dependencies, for which they propose a method.

While Beck et al. say '[w]e were able to replicate Oneal and Russett's (1997) original estimates exactly' (1998: 1276), they add, '[a] different picture emerges, however, when we correct for temporally dependent observations using grouped duration methods' (Beck et al. 1998: 1276). Beck et al. find that the dramatic effects are on the 'trade' variable. Beck et al. (1998: section 4) show how the data can be reanalysed to remedy the problem, though, as they say, the results might be very different.[18] They show that alternative method 'upholds' the effects of democracy but not the effects of liberal trade policies (1998: 1278).

An interesting example of progress here is Beck et al.'s exact replication of the DP results using the same methods, and the acknowledgement that even when the corrections are made the powerful effects of democracy remain intact. Furthermore, in subsequent papers Oneal and Russett (1999a: 22 and 2001: 470) acknowledge the criticisms and revise their analyses to meet them. Referring later to their 1997 paper, Oneal and Russett say, 'we did not consider whether there was heteroskedacticity in the error terms, account for the grouping of our data by dyads, or address the lack of independence in the time series' and add that Beck et al. 'showed us the error of our ways' (2001: 470).

Russett and Oneal manifestly take the criticisms seriously and address them in a way that, presumably, the critics would applaud. Oneal and Russett (1999a) correct 'for temporal dependence' in the estimation of their liberal-peace hypothesis equation. The serial correlation in the time series could lead to the coefficients appearing to have lower standard errors, which would erroneously elevate the measure of statistical significance. To deal with the problem Oneal and Russett 'report … the results of two analyses in which corrections for temporal dependence were applied in the estimation' of their liberal-peace equation (1999a: 222). In accord with the recommendations of Beck, Katz and Tucker, Oneal and Russett estimate the coefficients of their 'theoretical variables in the presence of a piecewise linear spline of the number of years since the last dispute' (1999a: 222). They contend that this and other adjustments they make to take account of recent methodological innovations provides further support for the liberal-peace hypothesis.

In the third example, Green et al.'s (2001) critique of pooling time series attacks a broad range of studies in the social sciences but they single out DP studies because of their prominence and importance. In pooled time-series analysis, dyads are put in an undifferentiated pool by year (or some other temporal measure). Such time-series pools are extremely common in studies of the DP hypotheses, by supporters and opponents alike. Since the published results clearly give supporters the upper hand (given the view that there is near-consensus), discrediting this sort of analysis would undermine their side more than the opponents. Proponents of DP hypotheses, particularly Russett and Oneal, have responded directly to these criticisms. With a revised 'fixed-effects' method, Green et al. test DP hypotheses and show that neither democracy nor mutual trade have strong pacific effects.

Russett and Oneal (2001: 311–12 and Oneal and Russett 2001) do not dispute this methodological argument entirely. They point out that dyadic analysis shows important effects that neither state-level nor systemic-level analysis is able to do (Oneal and Russett 2001: 483). They say that '[a] fixed effects model loses a great deal of information when used to explain militarized disputes … because many dyads have not experienced any armed conflict. Consequently, their experience does not enter into the estimation' (Russett and Oneal 2001: 312). They also cite others' support for their methodological position, like Beck and Katz (2001), who have been critical of some DP studies, as noted. But most importantly from the point of view of cumulation and approach-to-consensus, they re-test their data using fixed-effects analysis. They show that the DP hypotheses are supported by such tests and that the reason for Green et al.'s different results, is that the latter use a much shorter time period. Oneal and Russett (2001: 471–2) argue that a study like Green et al.'s, which covers just forty-two years, thirty-nine of which are during the Cold War, is likely to overlook much that an examination of 107 years will find (like Oneal and Russett's), because of the greater variation of the values of the variables in the longer period.[19]

Two important points are clear from the dialogue between defenders and critics of DP hypotheses. First, the DP defenders do not dismiss the attacks. They often take the methodological and theoretical criticisms very seriously and respond to them in the most direct possible way. They do not just shift indicators, terms and definitions to produce better statistical results but in most cases they develop definitions that capture more of the substance of the concepts at issue. Second, defenders have adhered to the core principles of the DP literature, that regime type explains conflict in that democracies (or 'liberal' or otherwise 'free' states) do not go to war against one another. They have not watered them down or shifted them so that they fail to capture the core concepts of DP. Thus they have not responded in a way that Lakatos would regard as degenerative (e.g., as Vasquez 1997, 1998, 2002 has argued in the case of neorealism). The work of Russett and his colleagues, which has been at the centre of the defence of DP hypotheses, has not abandoned the hard core of the Kantian notion of 'liberal peace', with its tripartite character of political freedom, economic interdependence and international institutions. While many corollaries have been tested, they have not been proffered as substitutes for discredited core DP hypotheses. Over the past decade and a half the work of Russett and Oneal has shifted from an emphasis on the political dyadic hypothesis to include also other elements of the Kantian theoretical prescription, economic interdependence and, more recently, international organisations. The latest book by Russett and Oneal (2001) specifically refers to all three factors in the subtitle.

The constructivist critique

Elements of the critique

Constructivists are generally critical of DP studies. The most sustained constructivist examination of DP studies is by Barkawi and Laffey (2001) and the contributors to their edited volume. The authors argue that events cannot be understood without historical context, which DP studies have overlooked as a result of adopting a definition of 'liberal democracy' that transcends historical epochs and conditions. Precisely which states can be regarded as democratic can only be determined in the context of the historical conditions prevailing at the time. The approach taken by DP supporters, 'with its transhistorical causal law based on fixed definitions of democracy and war and a nation-state ontology of the international, is far too simplistic a frame from which to analyse the various historical and contemporary configurations of democracy, liberalism, and war' (Barkawi and Laffey 2001: 2). In this connection Barkawi and Laffey raise the problem of the war of 1812 between the world's two leading liberal states. They point out that, conveniently for DP proponents, the nearly universally used correlates-of-war data begin in 1815. Russett argues that Britain did not qualify, on grounds that 'it just did not fit the criteria either of suffrage or of fully responsible executive' (Russett 1993: 16).

Barkawi and Laffey (2001: 13) say, 'A fixed definition of democracy or liberalism will reflect a particular historical moment, a particular set of social, political economic and social [sic.] circumstances; and a particular understanding of what being democratic or liberal entails.' There should not be a fixed notion. They point out (2001: 13) that 'Dryzek (1996) argues that democracy should be understood as a project' because '[d]emocracy means different things in different times and places' (2001: 13). They add that the decision regarding what counts as a democracy 'is inherently a *political* decision' (2001: 14).

Constructivists charge that existing empirical studies begin with an assumption about the presence of a causal relation and the direction of causality. Barkawi and Laffey say that the constructivist critique, presumably unlike the 'rationalist' realist and liberal debate over DP hypothesis, refuses 'to assume in advance the most significant causal relations, their directions, or their consequences' (2001: 4–5). Constructivists hold that scholars must not overlook the effects of globalisation and the internationalisation of capital (Jenkins 1987). Furthermore, mainstream studies of DP use an inappropriate ontology (Cumings 2001), especially as it 'presupposes classical liberalism's ontology of abstract individualism' (Rupert 2001: 153). They question also whether the politics of war and peace can be adequately understood when the only units considered are territorially defined states.

Constructivist critics assert also that the DP proponents' attributing peaceful relations among states to democracy or liberal attitudes 'betrays a deeply unreflexive attitude toward analysis' (Barkawi and Laffey 2001: 19). Mark Rupert adds that it 'is ideological in the dual sense that it obscures social relations and processes from the view of their participants, and by disabling critical analysis it implicitly promotes the interests of groups privileged by those relations' (Rupert 2001: 154). They thus seem to suggest that there is a self-congratulatory element in DP studies in that they praise their own liberal societies. Mann (2001) adds that such self-satisfaction is misplaced, since ethnic cleansing has been perpetrated by democratic and quasi-democratic regimes, which liberal DP proponents fail to notice because they misread the history of their own states. In a similar vein, Rupert (2001) questions the democratic nature of the market economies that liberal states develop both at home and abroad. Polities regarded as democratic have histories of internal violence in 'labor and gender relations' (Barkawi and Laffey 2001: 16), which undercut the claims about the peaceful internal norms in these societies. Furthermore, these societies often became democratic 'during the creation of a global system of empires, forged and maintained by colonial wars' (Barkawi and Laffey 2001: 17).

Evaluating the critique

Three comments are in order here regarding constructivist criticisms, dealing with their accuracy, with what they entail about the progressiveness of DP studies, and with the effect they have on the relationship to empirical DP studies. With regard to the accuracy of the characterisation of DP studies, Barkawi and Laffey seem to be off the mark on at least six points: the questions of acceptance of a fixed definitions of 'democracy', 'war' and 'state'; sensitivity to historical context; the use of the example of the war of 1812; the direction of causality; the interests of privileged groups; and, in at least one sense, the ontology of DP studies. With regard to accuracy, many of the issues that Barkawi and Laffey (2001: 19–20) say have not been disputed *have* been disputed: the meaning of 'democracy', the definition of 'inter-state war' and of 'sovereign territorial state' have all been raised, contrary to what Barkawi and Laffey say (2001: 4). The two preceding sections showed the scrutiny and analysis of the correlates of war and polity II and III definitions of 'inter-state war', MIDs and other sorts of conflicts short of war.

On the issue of historical sensitivity of DP studies, the critics are right that many of the studies adopt trans-historical definitions of key terms. However, it is incorrect to characterise the DP debate as a whole as doing this and of ignoring constructivist concerns. One of the widely cited authors in DP debate, Owen, argues (1994, 1997) that perceptions should play a leading role in explaining DP hypotheses, particularly as one formulates a plausible causal mechanism. Although Barkawi and Laffey cite Owen's work, they do not draw out the answers that it offers for some of their criticisms, such as Owen's view of the changing definitions of 'liberal democracy' and his treatment of some of the problematical cases. Owen's process-tracing applies a causal mechanism that involves the state's citizens' perceptions of potential rival state's regime type. Owen's position at least partially blunts the criticism based on the historical observation that the understanding of what constitutes a democratic or a liberal state may vary over time. 'What a scholar in 1994 considers democratic is not always what a statesman in 1894 considered democratic' (Owen 1994: 120).

Owen (1994: 102) offers a definition of 'liberal democracy'. But according to his theory, joint satisfaction of the definition may not prevent two democracies from going to war: the citizens on each side must regard the other side as a liberal democracy. Owen's inclusion of this perceptual element in the explanatory model allows him to explain the war of 1812 and the American Civil War. Russett also anticipates these constructivist objections in his explanation of why the Philippines war of 1899 should not be regarded as a militarised inter-state dispute between democracies. The perceptions at play 'at the time' were, he acknowledges, 'Western ethnocentric attitudes' (1993: 17). Though not himself subscribing to that approach, Russett is well aware of constructivist views and points out (1995: 165) that the postmodern constructivist 'would have serious reservations about the ability of an observer to penetrate the self-justificatory and mythological functions of decision-makers' texts to discern "real" motivations'.

The DP debate has not taken causal questions for granted, but has vigorously contested them. A decade ago Maoz and Russett (1993) tested the two most popular causal models (structural and normative) of the effects of democracy. There remains a serious debate about the causal processes (see Chan 1997: 85 and Layne 1994: 45). And there have been several statistical studies that specifically question whether democracy causes peace or peace causes democracy (see Thompson 1996 and Midlarsky 1992, 1995). Russett and Oneal (2001: 312) say, 'as we have frequently noted, it is very likely that there are important reciprocal relations between the Kantian variables and peace'.

Constructivists and postmodernists charge that DP studies have taken hold because the conclusions serve the interests of privileged groups (a charge similar to one they make against political realism). However, such a charge ignores the history of the DP debate. The most prominent IR figures who first commented on DP hypotheses (as advanced by Babst) were very hostile to it, especially Small and Singer (1976), who sought to 'explain away' the low incidence of conflict by invoking the proportionally smaller number of states that were democratic. It was only after Small and Singer's empirical tests were evaluated by others that strong support for DP hypotheses began to grow. Several mainstream IR scholars (Haas 1995, Kegley and Hermann 1995) specifically raised concerns about the dangerous policy implications of DP conclusions.

On the question of inappropriate ontology, this study argues that the proper conclusions result from an examination of the best theories. Chapter 2 argued that ontology should be developed only after considering an empirical evaluation of contending theories, and Chapters 2, 4 and 5 argue that best theory is selected on the basis of rational theoretical and empirical criteria. If it turns out that DP theories provide the best account of the phenomena under consideration, then that conclusion generates support for the ontology of those theories. In this connection, constructivists do raise important questions concerning the choice of what is to be explained. Why focus on peace between nation states and not on economic justice, racial and ethnic equality, or anti-imperialism? Constructivists are right to insist that those questions should not be crowded out by exclusive emphasis on peace between liberal democracies.

The second general comment deals with the effect of constructivist criticisms of DP studies, however well founded they may be, on the claim made in this section, that DP studies are progressive. For example, as noted, Barkawi and Laffey (2001) question whether a factor like 'democracy' could possibly have the same set of effects in radically divergent historical contexts. But consider that one scholar might publish an argument concurring with constructivists (a) that historical conditions change, and go on to argue (b) that the conditions of the post-Napoleonic era evolved only moderately. He or she could then argue (c) that liberal democracies do not go to war against one another. The debate between supporters and critics of DP hypotheses over the past quarter-century could be re-examined to see if there is progress on this (post-1815) subset of all of the available historical cases. Disagreements between scholars that were evident in the early 1980s would thus be diminished.

The third general comment deals with how much to support the constructivist approach over that of DP studies. This book acknowledges that constructivism adds a great deal to the study of IR. For example, how one group or individual understands or constructs the image of another group and how the process can be affected are certainly important in IR. But there are limits to constructivism, as there are to all approaches. Constructivists who argue that no other framework of study has value are wrong. As Chapter 5 argued, when constructivists deny the possibility of methods that allow for any sort of rationally, evidentiary- or theory-based conjecture about future consequences of present actions (i.e., prediction), they are wrong and they diminish any value IR may have for policy-making.

There is evidence of progress even in the area of the non-core hypotheses of DP studies. Consider the work of Kegley and Hermann (1995, 1997), who have examined the hypothesis that democracies are no more likely to intervene than non-democracies. Their first paper on this question (1995) argued that democratic states are more likely to intervene than non-democratic states. But after criticisms of their analysis, they revised their methods (e.g., Kegley and Hermann 1997) and found that their conclusion was incorrect (see Russett and Oneal 2001: 63 n. 14). Although this is a corollary of the DP hypotheses, it is another clear example of cumulation. Studies by Dixon (1993), Hewitt and Wilkenfeld (1996), Raymond (1994, 1996) and Rousseau et al. (1996), look at different facets of democracy and peace, but yield results that are complementary, and which thus produces a coherent picture. Rousseau et al. (1996) found that democracies are less likely to be involved in crisis initiation because they are more satisfied with the status quo. Hewitt and Wilkenfeld (1996) showed that the more democracies involved in a crisis, the less severe the violence that follows. Dixon (1993) found that international organisations' role in resolving conflict was much the same with democracies and non-democracies. And Raymond (1994, 1996) found that democracies are more likely to accept binding third-party arbitration but the outcomes were not more successful for democracies. The picture that emerges from these studies, as Chan (1997: 68) summarises them, is that 'as status quo powers, democracies are typically less likely to initiate crises. Once in a crisis, however, they do not preclude the use of violence, although they tend to limit its severity.'

The role of conventions

The conventionalist concept of the 'measure-stipulation' helps explain not only the broad pattern of dissensus in the history of IR but also helps explain the approach-to-consensus, where that has occurred in IR, such as in the DP literature. Two important conventional decisions that aid progress in the debate were noted above (p. 192). With those conventions, the consensus regarding the results of empirical tests holds up even when controlling for 'geographic proximity, economic development, and alliance membership' (Chan 1997: 64). Authors central to the debate on the DP hypotheses acknowledge the importance of the conventional element. There are reasons, moreover, why war between sovereign states is taken as the core dependent variable. Again, this is by convention. Oneal and Russett (1999b: 2 n. 2) say, '[b]y convention in the social science literature, war is defined as a conflict between two recognized sovereign members of the international system that results in at least one thousand battle deaths'.

Most conventions, as has been stressed throughout, are not purely arbitrary. Some possible conventional choices change or undercut the substance of the hypothesis. Consider the question of *when* the regime type is examined for each state in the dataset tested. One might adopt the convention (as is done by Kegley and Hermann 1997) of, first, counting the regime type of every country on 1 January of each year, then looking for dyadic wars each year and then checking to see how may dyads at war were comprised of two democracies and how many were not. However, this procedure would code the 1974 Cyprus war as one between two democracies, even though it was not. The democratic government of Cyprus was overthrown on 16 July 1974, prior to the invasion and was not restored until 14 February 1975, after the invasion. Thus a preferable convention would be to count the regime type on 1 January, if the state did not engage in war that year, and count regime type at onset of violence, if the state did engage in war.

One might also ask the reverse question, about the timing of the creation (rather than overthrow) of a democratic regime. That is, if a state accepts a democratic constitution, which imposes democratic institutions, on a given day and invades a democratic state three months later, would the attack count against the DP hypothesis? Some authors who stress democratic norms over structure have argued that it would not, since only 'well-entrenched' democracies have the opportunity to develop the norms of non-violence to a sufficient degree to behave as the DP hypothesis states (Russett 1993: 86–7).

Given that democratic structures lead to democratic and non-violent norms over a period of time, as norms take hold, how long a lag should be expected? Empirical study helps understand the time required in different societies for the norms to take hold. However, if 'five years' is chosen over five and a half or six years, the choice may, to that limited extent, be arbitrary.[20] Ultimately, it is a conventional choice how long one stipulates democratic institutions must be in place before the norms can be expected to take hold and affect foreign policy actions. However, it is not a completely arbitrary conventional choice, since sixty days is clearly too short and sixty years is clearly too long.

The purpose of this section is to show the convergence among researchers on the DP propositions, convergence that is interesting in light of the fact that the DP hypotheses do not fit comfortably into the dominant theoretical framework in IR, neither in terms of levels of analysis (where systemic theory dominates) nor in terms of substance (where realism dominates). It is not important for present purposes that any particular hypothesis (monadic quantitative, dyadic quantitative, dyadic qualitative) is proved or disproved. What is important here is the movement toward consensus and cumulation; the use of one scholar's results by another scholar who may not have previously agreed; and the increasing agreement on results, whatever they might be, which arise from increasing agreement on the appropriate methods, definitions and measures, which themselves require the stipulation of conventions.[21]

There is still much work to be done in the area of DP studies. Related hypotheses and corollaries need to be tested against different sets of evidence and causal mechanisms need to be solidified. There is still considerable dispute over which mechanism(s) account for the dyadic hypothesis. Such investigations help to show the reach and the limitations of liberal or democratic polities as pacific. And as a fallibilist account of enquiry, CC acknowledges that future theories (with a yet-to-be-proposed variable) or future evidence (a rash of wars between democracies) may one day force scholars to reject the dyadic hypothesis, however well confirmed it is at present.[22]

Conclusions

The CC view of empirical enquiry, and especially the social sciences, allows one to fulfil one of the goals of the 'new philosophy of social science' which is to explain actual patterns and practice of theory and enquiry. It allows one to answer questions, like i–iii posed above at the outset of this chapter, dealing with progress and approach-to-consensus of different fields. It has been argued here that IP, RU and HT fail on one or more of those questions. The conventionalist's concept of the 'measure-stipulation' helps explain not only the broad pattern of dissensus in the history of IR but also helps explain the approach-to-consensus, where it has occurred, most notably in the DP literature.

Social objects can be seen, in the non-epistemic sense of 'seeing'. But the access one has to social objects and processes is different from the access one has to physical objects, a difference that has significant consequences for the ease and likelihood of producing discipline-wide conventions on analogues to the measure-stipulation.

This book has argued that facing up to the conventional nature of social science theory can shed light on the state of dissensus IR theory, especially when the sort of access to social entia is borne in mind. The reason for dissensus in the social sciences, as compared to physical sciences or other natural sciences, has to do not with the conventional nature of theory, since both natural theory and social theory are conventional, but with the nature of the conventions, e.g., methods of measurement that theories in each realm require.

The present chapter has shown that CS, along with an account of the nature of the conventions used in various fields of enquiry, which stem from the sort of access to the objects of measurement, account both for the slow progress of IR and other social sciences and the much more rapid progress of the physical sciences, which neither IP nor RU are capable of explaining. The chapter also has shown that IP and RU lead to pessimism about approach-to-consensus, relativism and scepticism in IR, while CS does not. The conventional nature of empirical knowledge, together with the problems of access to the objects of knowledge in IR, does constitute a barrier to consensus, since theories are underdetermined by the data and there are always actual or possible unrefuted rivals. But it is a barrier that can, in principle, be surmounted, especially if discipline-wide ‘measure conventions’ can be developed, as in DP studies. Many economists would argue that their discipline has overcome the barriers – a development that neither IP nor RU nor especially HT can explain.

Lakatos’s account of scientific theory includes the fruitfulness criterion of ‘the prediction of novel facts’. If the terms ‘prediction’ and ‘facts’ must be interpreted along the lines of Lakatos, it is clear that causal conventionalism satisfies this criterion. The argument of this chapter, on pp. 180–192 above, is similar to Lakatos’s discussion of the history of physics. Consider his own example.

> Einstein’s theory is better than … Newton’s because it explained everything that Newton’s had successfully explained, and also it explained to some extent some known anomalies and, in addition, forbade events like transmission of light along straight lines near large masses about which Newton’s theory had said nothing but which had been permitted by other well-corroborated scientific theories of the day; moreover, *at least some* of the unexpected excess Einsteinian content was in fact corroborated.
>
> (Lakatos 1978: 44)

Similarly, CC explains what IP, RU and HT cannot explain, namely, success in the physical sciences, which IP and RU cannot explain, and successful approach-to-consensus in the social sciences, which the HT cannot explain. CC, along with a straightforward account of the differences in ‘access’ to physical and social objects, allows answers to questions i and ii above, concerning why approach-to-consensus has been so difficult in IR but much less so in physics.

While terms may be differentially defined by theorists in competing approaches in IR, it does not follow that paradigms are incommensurable or that there is no theory-neutral language capable of falsifying propositions. If T1 and T2 use different notions of ‘power’, it is still possible to propose a theory T3 that incorporates both notions, so long as the definitions are precise enough to avoid confusion.

The picture of IR as conventional obviates the need to choose between the claims that IR is scientific (naturalism), on the one hand, and that it is inherently dependent upon 'non-scientific' or extra-theoretical assumptions, on the other hand. It may well be both. One may acknowledge the conventional element in IR without any admission of an 'unscientific nature'. A conventionalist view of IR theory permits retention of a robust naturalism and at the same time provides an understanding of the key differences between the history of the social and the physical sciences. Just as the physical scientist must make a conventional decision about practical matters before concluding whether or not Euclidean geometry applies to modern physics, the IR scholar must make conventional decisions about methods and metrics before settling on acceptance of neorealism, liberal institutionalism or other theories of international politics.

Appendix: Carnap's solution to the dilemma of non-Euclidean geometry

Carnap offers one of the most innovative and powerful solutions to the dilemma for physical theory posed by Einstein's adoption of non-Euclidean geometry. He distinguishes different *sorts* of space: *intuitive*, *physical* and *formal*. Carnap maintains that much of the disagreement in the foundations of geometry emanates from the different senses of 'space' used unwittingly by different theorists. The distinction of 'formal space' from the others is the most straightforward. It is purely abstract and applies to undetermined objects, while (as with the others) the order relations are entirely determinate. The distinction between intuitive and physical space is more subtle.

We experience the physical world through our spatial intuition; 'essential insight' (as the term was used by Husserl, 1913) allows us to grasp the general features of this intuition, which we can order *a priori* within the Euclidean system. Physical space then orders in a precise mathematical form the objects of our actual experience; the latter occur in our *a priori*-constructed intuitive space. Physical space arises only when we subordinate actually-experienced phenomena to the synthetic *a priori* space of intuition.[23] Carnap in this way uses the notion of 'intuitive space' to link our physical theory to experience of the world that we 'immediately grasp', which are mediated through given synthetic *a priori* truths.

Carnap makes use of the concept of 'essential insight' to explain how it is that we reach synthetic *a priori* truths of geometry. Pure logic allows us to draw general conclusions that apply to all concepts and objects. Our understanding of other relationships holds only for particular classes of object. We are able to draw valid conclusions about relationships between colours (even if we are not experiencing them but only imagining them). Likewise, we can draw conclusions about relations between lines or planes without experiencing them. Euclidean geometry is presented to us through intuition, as we can use 'essential insight' to grasp the truth of Euclidean geometry.

How can an essential insight lead us to (erstwhile synthetic *a priori*) propositions of Euclidean geometry that are apparently untrue? The distinction between intuitive, formal and physical space helps answer this question. On Carnap's account the geometry we are given as true, which is synthetic *a priori*, is the class of all possible Riemannian manifolds. But we work with specific geometric systems: the space we experience is Euclidean, and the space of relativity theory is non-Euclidean. Carnap maintains this by distinguishing the smallest regions of experience – 'infinitesimal space' – from the much larger regions of 'local space', and both from the largest of regions, 'global space'. The former is Euclidean, while the latter two are not. In other words, the axiom system that includes the parallel postulate does indeed hold for intuitive space. Euclidean geometry is valid in this realm. Thus the propositions are *a priori* truths.

Carnap's is one of the most interesting and complete accounts because it not only salvages the fundamental core of Kant's position, but also helps explain how so many other geometers, philosophers and scientists (who distinguished 'applied' or 'empirical' from 'pure' geometry), and who so strongly disagreed with Kant, also managed to disagree with one another.

7 Conclusions

> This is the bitterest pain to human beings: to know much and control nothing.
>
> Herodotus, *The History*, IX.16, tr, David Grene

IR theory is capable of fulfilling the historical mission of the discipline – of aiding policy-makers by providing a rational basis for the attempt to control future outcomes. Policy-makers require that theories, in order to be of any use, must be able to generate predictions, even if they are only probabilistic and approximate. The CC account of IR theory defended here includes criteria to help guide one's choice among predictive theories.

The preceding chapters discussed the accounts offered by influential philosophers and social scientists who defend reflectivism, SR and critical realism, as well as traditional naturalism with its scientific approach. The CC account shares some common ground with many of them, though differs substantially from all. A comparison of the CC account to the views endorsed by contemporary IR authors shows that there is perhaps the most common ground with that of the scientific approach taken by Vasquez, whose meta-theory stresses the attempt to discover extra-logical but rationally justifiable and non-arbitrary grounds for theory choice. Vasquez's Lakatosian methods are not, however, fully embraced here. In the reflectivist tradition, Bernstein et al. offer a method of analysis which is more grounded in naturalist meta-theory than they seem to acknowledge, but which provides a means of applying the meta-theoretical conclusions reached in Chapter 6 about prediction. While the scientific realist views of Wendt are not accepted, many of other elements of his view are defended here – although often on different grounds. One example is the importance Wendt places on both capabilities and ideas in IR theory, which need not set them up as mutually exclusive in theory, necessitating an 'either/or' choice.

CC shares elements of critical realist authors, especially their emphasis on the role of causal hypotheses. There are, though, many points of divergence. Patomäki repeatedly says that 'social structures are real'. But it has been argued that there are no clear benefits from such a claim, either in theory or in the application of theory to policy-making. Patomäki infers the reality of social entia only from their observed effects. Conventionalism, however, does not deny the reality of the world beyond sense perception.[1] In the philosophy of social science, the work of Little (1991) and Bohman (1993) are cited (especially in Chapter 5) as providing valuable insights from causal realism and hermeneutics, respectively. Work in the philosophy of the physical sciences, especially Duhem's under-utilised principle of CS, informs much of the theoretical direction of this study.

The argument of Chapters 2–6 rejects the HT doctrine of the unavoidable subjectivity of social enquiry, especially as developed in postmodern frameworks, and rejects the principles of RU and IP. The study has offered a way to defend the policy-value of IR by means of the CC account. CC, as it is developed here, includes

- defences of CS
- Duhem's limited underdetermination principle
- the distinction between 'observation' and 'theory'
- intersubjectively valid knowledge
- the possibility of rational and non-arbitrary criteria of theory choice
- a non-sceptical view of HT
- common-sense, motivational and causal realism
- a prediction capability for social science theory
- a role for norms in the study of IR
- reasons as a form of cause in the social sciences
- multiple theories and methodological pluralism and
- the possibility of cumulation and approach-to-consensus.

The principle of the conventionality of all science

The use of conventions is always a part of selecting a theory of empirical phenomena. The physical sciences, which involve conventional choices, are taken as the best and most complete of systems of empirical knowledge. Any claim that other disciplines are capable of more than the physical sciences cannot be assumed and must be supported by argument. This study acknowledges that because scientific knowledge of the physical world is inherently conventional, the same should be expected of knowledge of IR. The CS principle helps to answer some of the most important meta-theoretical questions and some of the puzzles about the history of dissensus and consensus in IR.

Use of the term 'conventionalism' invites confusion and criticism. However, the standard attacks are generally not valid against the Duhemian version advocated here. For example, a number of philosophers of science, led by Popper (1968), charge that conventionalism is unable to utilise empirical evidence to count against a theory or its laws. It should be clear from the version of conventionalism defended above, which may more accurately be termed, 'quasi-Duhemianism', that the doctrine is thoroughly fallibilist (see Chernoff 2002). The more well-known version of conventionalist philosophy of science developed by Poincaré depicts the reliance on convention as a way of insulating the theory from empirical rejection. If the laws of the theory are analytic and derive their truth from the meanings of the terms and logical connectives, then empirical evidence cannot disprove them. The version of conventionalism defended here follows Duhem (1954, 1969) and holds that good laws are true and bad laws are false. Popper (1968: 79–81) oversimplifies by conflating the views of Poincaré and Duhem on this point.

Moderated underdetermination of theory by data

The CC account here argues that there is a limited form of underdetermination of theory by data: the data will always be consistent with more than one possible theory. However, the stronger, radical principle of underdetermination holds that adjustments may be made so that any theory can be rendered consistent with any set of observations. The more moderate Duhemian underdetermination principle rejects this. Enough counter-instances will require rejection of a theory (and its system of auxiliary hypotheses). While there may be many theories consistent with the observation, CC holds that rational rejection of 'falsified' theories is warranted when enough falsifying evidence accumulates and when there are other theories which do not suffer the same defect.

Duhemian conventionalism is closely associated with the view that there are no crucial experiments in physics, which suggests that Duhem does reject 'falsifiability'.[2] However, Duhem's rejection of crucial experiments is not based on a rejection of 'falsifiability' but rather on the following argument: given theories T1, T2 and T3, each of which would yield a unique prediction (x, y and z) under conditions C, one would be tempted to conclude that if x occurs, T1 is true and thus T2 and T3 are false. But Duhem argues that there are no crucial experiments in the sense that the observation of x in the experimental circumstances C warrants neither acceptance of T1 as an incontrovertibly true theory nor the final rejection of T2 or T3.

As far as the conclusion that T1 is an incontrovertibly true theory, Duhem points out that there may be other theories not included in the list, which would, similarly, predict the outcome of the experiment correctly and would also predict other experiments that neither T1, T2 nor T3 could. So if x obtains in the experimental conditions, T1 might not be true, because there might be some other theory, T4, which also predicts x and which does better than T1 in other situations (perhaps not yet encountered). Theories T2 and T3, which allegedly have been 'disproved', cannot be rejected with certainty because the connection between T2 and y and between T3 and z require auxiliary assumptions (e.g., about the functioning of the measuring equipment, etc.). It may turn out that some of the assumptions are false; correcting them may enable T2 or T3 to yield correct predictions.

While the natural science analogy breaks down at points and the social sciences are, in those respects, at a disadvantage in achieving properly 'scientific' knowledge, the social sciences also have some advantages over the natural sciences. The natural sciences – molecular biology, astronomy, physics, etc. – move forward by demanding specific observations to help theory choice. These observations often require the construction of new mechanisms and experimental apparatuses that previous generations of researchers were denied, be they cathode-ray discharge tubes, x-ray diffraction of proteins, or telescopes. However, the interpretation of the perceptual observations to produce an 'observation' relevant for theory choice requires that the researcher accept an often highly elaborate theory of how the apparati themselves operate and produce different (observable) effects that can be applied to problems of theory choice. As Duhem shows, all theory choice requires accepting (and rejecting) not just an hypothesis or theory but a complex of theories-plus-hypotheses about experimental procedures-plus-auxiliary hypotheses. This creates the possibility that the theory choice between T1 and T2 is hampered by the researcher's or scientific community's erroneous acceptance of another theory regarding the physical (or chemical) operation of the experimental apparati. IR theory is subject to other problems of reflexivity, interpretation and hermeneutic complexity that the natural sciences are not, but it is largely free from this additional obstacle.

Theory, observation and norms

The traditional distinction between 'theory' and 'observation' has been heavily criticised over the past thirty years. There is much merit to the arguments against some methods of drawing the distinction, especially those found in some classical empiricist writings. However, as Chapter 3 showed, 'theory' and 'observation' are unfairly confused by the failure to separate other distinct concepts. Dretske's argument distinguishes 'epistemic seeing' from 'non-epistemic seeing' and permits identification of a genuine sense of 'see' or 'observe' that does not require that it be embedded in a theory. Chapter 6 recounts Greenwood's analysis showing that, even if one rejects the distinction between 'observation' and 'theory', the inferences often drawn to other principles like IP, which lead to further scepticism, are unwarranted.

With respect to norms, even if one denies a hard-and-fast 'fact–value' distinction, Chapter 1 showed that the value-content of statements differs greatly. This study has endorsed roles for both norms of reason and norms of action. Although the CC account is naturalist and allows theories to be descriptive, explanatory and predictive, it leaves open a role for moral and evaluative theories. Chapter 1 showed that moral principles and precepts require description and prediction in theories because one's moral aims cannot be fulfilled if one does not know what actions will lead causally to their being brought about. This view follows for consequentialist and deontic theories alike. One must know what is possible and probable in order to make or appraise foreign-policy decisions. In international politics one may not blame a leader or a state for failing to achieve a certain goal if there was not the means or capability to bring it about. Policymakers must have theories that are primarily moral and theories that are primarily empirical.

The possibility of non-arbitrary, rational criteria of theory choice

This study has defended a set of rational criteria of theory choice. There are ways of selecting one theory over another that are not arbitrary and can be defended by use of rational, philosophical argument. Some of the criteria are logical, such as internal and external consistency. An internally consistent (i.e., self-consistent) theory is to be preferred over an inconsistent rival even without examining the data or the relationship of the data to the theory. The same is true for externally consistent theory (i.e., one consistent with observations) since it also, ultimately, avoids the acceptance of contradictory propositions. If the theory entails x in all circumstances C and there is an observation of not-x in circumstances C, then the theory is held to deficient.[3] CC holds also that there are rational, non-logical criteria of theory choice. Popper's criterion of falsifiability is an example. Although it is *non-logical*, it is widely accepted by IR authors and even invoked by some postmodern theorists. On this criterion, a theory (when conjoined with other components of its research programme) for which there are conceivable falsifying conditions is to be preferred over one for which there are none. This criterion is non-logical in that it is not based on a logical principle, like *principium contractionis* or *modus tollens*, in the way that the two logical criteria just cited are.[4]

It is possible for one to criticise conventionalism by highlighting the assumption of 'rational enquiry', noted in Chapter 1, upon which it relies. The conventionalist position advocated here argues that there are rational grounds for theory-acceptance that are found outside of deductive logic and outside of any given theory. If an interlocutor claims not to be rational, does not endorse the canons of rational discourse, and does not base any actions on such canons, it is difficult to carry on any further dialogue. Agents who genuinely reject rationality are likely to come to grief sooner (by disbelieving in rational grounds for the claim that it is dangerous to step in front of a speeding train) or later (disbelieving in the dangers of repeatedly leading a militarily weak state into battle against a powerful coalition). Argument alone is not likely to sway interlocutors who insist that they reject any value that rational or philosophical enquiry might have. Thus, a critic does have the option of generally eschewing rational grounds for foundations of IR and foreign-policy-making (and thus of ignoring the preceding six chapters). The conventionalist position, along with most others, begins with the assumption that enquiry is rational, namely, that philosophical enquiry has value and allows one to arrive at truths. Such truths form one step in the reasoning that permits the foreign-policy-maker to achieve desired goals. Erroneous philosophical and meta-theoretical principles will produce incorrect criteria of choice, which in turn will lead to the adoption of inferior theories, which in turn will lead to confidence in false laws. On the basis of such laws policies will be chosen in the mistaken belief that they have the best chances of achieving diplomatic, economic or security goals for the state. If a critic persists in arguing that the sort of rational criteria that CC endorses are irrelevant, the only retort is to note that if philosophical foundations are abandoned in the debate over conventionalist theories of the philosophy of science and social science, then they must be abandoned everywhere.[5]

The social sciences observe the Duhemian norm of rational enquiry and theory choice. If one's work is part of a genuine rationally based discipline, then the community will recognise when the evidence supports one approach and leaves the existing alternatives behind (though there may be as-yet-unimagined alternatives that do better). This norm is found in IR literature, like DP studies discussed in Chapter 6. Russett and Oneal say:

> [s]upport for the Kantian peace, especially the benefits of democracy and economic interdependence, is extremely robust, as we have tried to show in this book. No variable is significant in every possible test, but if the weight of evidence is considered, there is little doubt that liberals were right: democracy, economic interdependence, and cooperation in international organisations reduce the incidence of war.
>
> (Russett and Oneal 2001: 313)

This is an example *par excellence* of the Duhemian confidence in the weight of evidence from a series of inquiries that leads a community of investigators to a conclusion.

Intersubjectively valid knowledge and a non-sceptical view of the HT

A fallibilist, pragmatist theory of knowledge regards statements about the empirical world as interconnected in ways that either mutually reinforce or contradict one another, leading to greater confidence in or rejection of specific claims. Some singular statements, laws, etc. are reinforced so well that they achieve high confidence levels. But they never achieve certainty. Still, such theories are non-sceptical. One might view the process of interpretation in the HT in a similar way; the circularity of interpretation leads to a similar dialogue between statement and interpretation. While 'certainty' about the best interpretation is not possible, there are good rational grounds for accepting one over others – in the process of what Peirce calls 'the fixation of belief'.

Since scepticism does not follow from fallibilism, why should it follow from HT? Bohman's theory was appropriate to discuss because Bohman is one of the HT authors who makes a powerful case against the scepticism usually regarded as inherent in the HT, although he remains sceptical of prediction. In the case of fallibilist theories of scientific knowledge, there is never a point where 'a final theory is finally proved'. There may always be other theories that no one has yet imagined that might fit with the existing evidence better, and there is always more evidence that could be gathered that might shift the balance from the currently accepted theory to a known rival. For this reason empirical knowledge may allow and enable successful practice in the real world but, as noted, it does not warrant certainty. Given the model of the fallibilist, pragmatic theory of knowledge that allows the warranted acceptance of propositions, it would seem that the HT should be able to develop interpretations that could warrant knowledge about description, explanation and, as Chapter 5 argues, prediction in a parallel way.

The CC account envisions a dialogical relationship between the interpretation of the statement of a question or problem (since the formulation of the problem will presuppose an ontology) and the theory that provides the answer (since the theory also implies an ontology). The theory and a new question each may be offered as a way of refining one another. A new question is posed against the background of an old set of theoretical beliefs; a new theory is formulated to answer that new question. The new theory, with its new ontology, may warrant a reinterpretation of the question or problem.

Bohman says that the 'new philosophy of social science' begins with actual scientific practice, and builds on 'successful social explanation', even though he (surprisingly and disappointingly) never defines, characterises or explicates that concept. The problem posed for Bohman by the study of democratic peace is that it would seem to be as good an example as one can find of 'successful practice'. But DP theories are powerful enough to generate predictions – at least general and probabilistic predictions about the long-run futures of democratic–democratic and democratic–non-democratic interactions. As noted above, the positive picture of explanation Bohman paints and his sceptical view of prediction result from unequal treatment and various biases discussed in chapter 5 above.

Many reflectivist critics of naturalism cite the limitations on 'objective' or 'intersubjectively valid' knowledge that arise from considering the hermeneutic nature of social enquiry. Bohman's account is very useful here because he argues successfully that most reflectivists overestimate the limitations on 'explanation' arising from the sorts of circularity that inhere in social investigation. He argues that developments in the philosophy of social science acknowledge the indeterminacy of social enquiry and have made 'it manageable within empirically adequate and verifiable explanations' (1993: 232). Bohman emphasises rational-comparative interpretation, described in Chapter 5. Sceptics and critics have ignored this sort of indeterminacy, focusing instead on 'holistic-contextual' interpretation, which leads them to an unwarranted degree of scepticism. However, Chapter 5 contends that Bohman's argument could be taken a step further, from 'explanation' to 'prediction'. The grounds for a notion of 'prediction' in the social sciences are parallel to those for 'explanation'. The HT could include an account that justifies limited forms of prediction.

This study acknowledges reflexivity of human action and the double rule-governed nature of the circularity of interpretation but disputes the sceptical conclusions that are often taken as flowing from them. On the question of interpretive circularity, this study holds that Bohman's 'profoundly anti-sceptical' arguments are compelling but that he does not take them far enough.

Common-sense, motivational and causal realism

The position advocated in this study rejects SR, but endorses a number of other philosophical doctrines that fit under the heading of 'realism'. One example is common-sense realism, according to which tables and chairs and professors and snow-covered country lanes exist more or less as we perceive them. With respect to the sciences, there is an underlying reality that our empirical researches move closer toward, and our belief in such a reality is a useful and proper motivation for some of us. The belief that there is an underlying reality which motivates much scientific research, but which any available theory is capable of capturing adequately, is known as 'motivational realism' and is also endorsed above. But this is very far from the scientific realist view that science's 'latest theory' embodies a correct picture of that reality, or that investigators should accept an ontology that embodies it.

This book attempts to show how to develop a minimal foundation for IR theory (and, more generally, for allied social sciences). It dispenses with unnecessary baggage that may when unpacked be found to be vermin-infested or otherwise unwholesome. What one may do without, one should do without. While the study of ontology and metaphysics is perfectly respectable, authors like Wendt, Patomäki and Dessler, who insist on an ontology that requires a form of reification of theoretical entities, can lead to dangers without offering anything particularly useful for the theorist or policy-maker.

Vasquez endorses the need for clear criteria of theory appraisal, whether for empirical or normative theories. And he argues that there are rational bases for such criteria, even if they cannot be deduced from strictly formal or logical premises (as logical positivists had hoped). However, he talks of 'true theories' and he talks of the possibility of falsifying specific theories or hypotheses outside of the holistic framework. He says, '[t]hus, while specific theories or explanations may be falsified …' (Vasquez 1998: 232). The arguments adduced by Duhem show quite persuasively that specific theories, and *a fortiori* explanations, cannot be falsified in isolation from the research programmes in which they are embedded, including beliefs about experimental procedures, equipment, etc. It should be emphasised that, while the CC account presented here does not recognise the 'truth' of theories, even the best theories, it does emphatically endorse the usual sense of 'truth' in IR discourse: there are, as Chapter 3 argued, true observations, true generalisations and true laws. But the CC account suggests suspending judgement on the truth of explanations or theories. Theories, and the explanations which in part constitute them, are appraised on grounds of superiority over rivals without imputing truth or falsity.

While the conventionalism defended here is conjoined with the endorsement of common sense realism, it need not be. One might very well accept conventionalism and reject both SR and common-sense realism, which allows one to accept observed entities into the ontology. Rejection of SR denies acceptance of theoretical entities, even though they may exist (see pp. 51–2 and Chernoff 2002). Observational causal laws have causal mechanisms that explain their effects and the entia are regarded as adequately demonstrated. Hypothesised theoretical entities show that there are plausible causal mechanisms associated with theoretical laws. However, there is no reason to regard any specific theoretical entia as 'real'. While a theoretical law should be accompanied by a plausible causal mechanism, on the CC view, that law entails the existence only of the observational entia implied by the observational law(s) it explains.

Prediction

Any theory that does not account for prediction is without empirical value for policy-makers, as Chapter 1 showed. One should seek a theory that allows for prediction, unless predictive theory can be proven illegitimate. The standards of proof for any anti-predictive argument must be very high. The Eleatics argue that change is an illusion, which conflicts with virtually all experience of the sensible world. Because of the enormous number and diversity of previously accepted truths that would have to be surrendered to accept this claim, one should demand a very high standard of proof for it. Experience does seem to support (non-point) predictions of human behaviour. For example, there seems to be little problem with predictions of the behaviour of individual humans such as: the hungry baby will cry some time during the night; or of states such as: France will not invade China in the coming year. Any theory that prohibits prediction will, like the metaphysics of Parmenides and Zeno, require an extraordinarily high standard of proof, because the alternative appears to be so well confirmed.

The examination of anti-predictive arguments drawn from a variety of sources (such as non-linearities, social complexity, the absence of governing regularities) showed that there is no conclusive argument against the possibility of predictive theory. And prediction indeed seems possible in international relations, albeit with certain qualifications. The foregoing has acknowledged qualifications on the predictiveness of social science theory. Predictions are probabilistic and their strength is limited by the value of observed empirical associations and by the future temporal frame (since they are less reliable as the time-frame is extended, which follows from the axioms of the probability calculus). However, the calculations produce better results than randomly chosen policies. And random policies are the alternative if one rejects belief in rational calculation and causation on which it is based.

The review of the attacks on prediction showed the arguments to be fundamentally flawed. Either they derive their conclusions by means of a straw man (an uncommonly narrow definition of 'prediction' that presupposes many unreasonable conditions) or the accounts supposedly inconsistent with prediction in fact allow, on closer inspection, room for prediction. For example, Chapter 5 found the scenario method defended by Bernstein et al. attractive, though for reasons not fully acknowledged by the authors. The method turns out to be an application of the predictive approach and not an alternative to it, as Bernstein et al. claim. In Chapter 5 (pp. 135–7) Bohman's HT account was shown to be consistent with a notion of 'prediction', as well.

Methodological pluralism and multiple perspectives

The subject matter of IR, like that of other social sciences, is complex and has many facets that can be approached from many angles. This book does not purport to say all that there is to say about meta-theory. It aims to comprehend some of the facets of the study of international politics and the social world. The elements of the CC account (observational versus theoretical laws, reason as causes, the CS thesis, probabilistic prediction, etc.) are defended because they are among the most important in a meta-theory; they are needed to answer the central questions of theory-building and theory choice discussed in Chapters 2–6. The book endorses a naturalist-oriented account, but since the study of IR involves vastly different sorts of questions, it accepts that IR theorists ask some questions that require a more purely interpretive approach and may not offer much to the policy-maker. This study does not rule out the pursuit of knowledge for its own sake. But it defends the study of IR against the charge that it is capable of yielding only that and nothing more.

The results defended in the foregoing are entirely consistent with the requirement of methodological pluralism as described in Chapter 1, which endorsed the need for different sorts of methods, both quantitative and qualitative, often drawn from techniques developed in other disciplines (natural sciences and formal sciences, humanities and other social sciences). Likewise, different research questions demand different sorts of theories, which operate at different levels of analysis and entail different ontological commitments. There is no way to construct a single, over-arching theory to do all the work that scholars legitimately demand of the discipline of IR. Cross-cutting theories may be used to solve related problems. Physicists use theories of quantum phenomena to answer some questions (e.g., to explain and predict the behaviour of some parts of their domain), use special relativity to answer other questions, and they even use Newtonian mechanics to answer still other questions. Theories that explain the actions of Andrei Gromyko or Woodrow Wilson are different from ones that provide policy guidance about national security if China and Saudi Arabia were to become liberal democracies; they are different both in terms of levels of analysis and the range of variables they consider.

Different kinds of theories are needed for different kinds of problems. While this study endorses a form of naturalism, it does not deny that alternative theoretical orientations are preferable when confronting certain sorts of problems. Why do wars occur in the international system? Why did Cyrus Vance resign as Secretary of State? Why did the UK stay out of the Eurozone? Why did the US invade Panama in 1989? Why were there few great-power wars between 1815 and 1914? Each of these questions involves a different level of specificity. The sorts of actors to be considered vary, as do the set of appropriate variables.

The view presented here contrasts with aspects of the HT, postmodernism and critical realism, as well as with older forms of behaviouralism in that these often permit only a single type of theory or method and banish many types of enquiry that do not conform to their strictures. For example, as Chapter 6 noted, constructivists Barkawi and Laffey (2001) argue against the data-based historical studies that DP authors typically use. Similarly, in stressing the stratification and levels, critical realism appears to be aiming at a unified social science theory which is not possible. Certainly not if there are distinct theories that operate at different levels of generality, and with different research methods (see Patomäki 2002: 81–5). Multiple perspectives and methodological pluralism preclude a 'unified' theory. CC argues for (and IP excludes) the possibility of theoretical convergence. But the convergence that is possible is among competing theories driven by the same research questions, that operate at the same level of generality. It is not convergence of anything like a social science version of a unified field theory.

Richard Little's (2000) interpretation holds that the English school was methodologically pluralistic. While others have argued that differences of methodology result from differences between different English school authors, such as Bull (1977), Butterfield (1966), Watson (1992) and Wight (1991), or from those authors changing their views over time, Little holds that differences between their ontologies of international systems versus international societies versus world communities account for the differences. For example, Little finds Bull's idea of 'international systems' to be rather close to Waltz's conception, and argues that systems are studied by what he calls 'positivist methods'. The present study finds much common ground with the English school's understanding of ontology and the appropriate methodologies for each type of object of study and question involving those objects.[6]

Scientific enquiry in international relations

This study acknowledges the need for moral concepts, principles and theories in foreign policy-making. It also shows, as one of its chief conclusions, that rational policy decisions are possible. Many other foundational positions deny this, often without acknowledging it. There are then grounds for a strong connection between what IR theorists do and the intellectual tools that foreign policy-makers use as they attempt to change the world.

The study emphasises the fallible nature of social-scientific knowledge. This is in no way a defect in social science but rather is a feature shared with all empirical knowledge. There are many limitations on social science knowledge that are not found in the natural sciences. However, theories and laws in either domain may be overthrown when the community of investigators uncovers overwhelming counter-evidence, as physicists and astronomers did with Aristotelian and Ptolemaic laws of planetary motion, and as IR researchers did with Babst's and Rummel's claim that democracies are inherently more peaceful than non-democracies.

This book set out to resolve the tension between the meta-theorists and philosophers of science who argue against the predictive ability of IR theory and the theorists and policy-makers who assert or make use of predictions. It has defended a justification of the use of prediction based on the view of the social sciences that has been termed 'causal conventionalism'. The naturalist argument of causal conventionalism in Chapters 2–6 shows that progress is possible in the study of IR. This is clear especially from the discussion, which argues that DP studies constitute an example of progress in IR theory; *ab esse ad posse*.

The CC account recognises an essential role for moral and normative discourse to work alongside behavioural theories. The CC account of IR theory thus offers a grounding for the rational and humane formulation of foreign policy that avoids the scepticism of IP and RU principles and of many HT theories. At the same time it recognises the role of the social science equivalent of the physical 'measure-stipulation', the means to evaluate theories rationally (albeit in holistic fashion). Together these create the possibility of much greater progress in cumulation and movement toward rationally grounded consensus positions in IR and the possibility of theoretically based predictions to support the formulation of foreign policy.

Notes

1 Policy-making, prediction and the theory of international behaviour

1 'Rational' here is meant broadly enough so as not to exclude morally based theories of foreign policy.
2 Still, reaction by theorists can be so powerful that a less loaded term may be preferable, such as 'rationally based expectations' or 'prognostication'. Doran (1999) uses 'point prediction' as a more precise forecast (see p.161). Nevertheless, the term 'prediction' is used in this study, since it is the term typically used to capture the notion defined on p. 8. Ray and Russett (1996) use the terms 'forecasting' and 'prediction' interchangeably. In contrast to 'prophecy', they define 'prediction' as 'a statement about the future based on specified contingencies' (Ray and Russett 1996: 467).
3 Popper supports a more limited form of justifiable inference about the future, which is involved in 'piecemeal engineering' (1971a: 157). His position is, then, consistent with the conclusions drawn in this book. Popper's distinction and his arguments are discussed further in Chapter 5.
4 The idea of non-theoretical knowledge is somewhat complex, since it may involve the use of theories in order to understand the meanings of the terms used to express the 'fact'.
5 See Wendt 1992, Suganami 1996, Suganami 2003 and English school distinctions among Kantian, Lockean and Hobbesian forms of anarchy.
6 Within the limits imposed by Duhem's argument. See pp. 69–70, 104–6 and 167–9.
7 In the revised edition (1998) of his important examination of political realism, Vasquez relies on the criteria of 1) accuracy, 2) falsifiability, 3) maximum explanatory power 4) Lakatos's progressiveness of research programmes, 5) consistency with knowledge in other areas and 6) parsimony (1998: 230).
8 One might be tempted to object that, while there are surely some constraints on social knowledge, they may not be the same as those on physical knowledge. Hence it is an unacceptably anachronistic form of naturalism, and thus a departure from the last several decades of debate in IR meta-theory, which would assume that any constraints on physical knowledge would apply to social knowledge. But this objection is not consistent with the very IR meta-theory debate to which it appeals. For one of the central elements of that debate is over Kuhn's IP thesis. And the debate proceeds on the assumption that the IP constraint that Kuhn attributed to the natural sciences (especially physics) applies also to IR. There is no line of argument within that debate over whether IR is subject to the constraints of IP *even if* it can be shown that physical science is constrained by it.
9 As Little sees the English school of Butterfield, Bull, Watson and Wight, positivist methods are appropriate for enquiries into questions of the international system,

interpretivist or hermeneutic methods for international societies. 'International systems and international societies, therefore, rest on very different ontological assumptions and, as a consequence, they need to be examined by means of very different methodologies' (Little 2000: 408). See below, pp. 216–7.

2 Social sciences, naturalism and scientific realism

1 Popper is one of the few twentieth-century philosophers to be seen as a major figure in two very different areas of Western philosophy, philosophy of science and political philosophy.
2 '… men do not think they know a thing till they … grasp its primary cause'. Aristotle (1956), *Physics*: 194b.
3 The point here is to note several influences on social science theory. This does not assure us that these criticisms of positivism are uniformly accurate. Graham Bird (1995) has argued that the fundamental assaults of Quine miss their mark.
4 This position is considered in more detail in Chapter 6. On the principle of radical underdetermination of theory by data, see Quine 1990: 95–105.
5 Kuhn applied his account only to natural science. He explicitly reserved it for 'mature sciences'. Others have extended that account to the social sciences.
6 See Putnam 1975 and Boyd 1973. Criticisms of SR are presented in Chernoff 2002.
7 See Ragin 1988. See Chernoff 1995 for an application to a question regarding the NATO alliance.
8 Quester discusses 'self-fulfilling prophecies' and argues that predictive theory is needed despite the problem of self-fulfilling prophecies (2002: 208–9). He links this problem to the relevance of an IR version of Heisenberg's uncertainty principle. He says (2002: vii–viii) that, 'any political process that is watched closely is changed in the process of being watched … any attempt to predict will see that prediction itself becomes a factor in the entire process'. Quester suggests that the predictive value of IR is closest to the physical sciences' weakest predictive discipline, meteorology. (Quester notes that he served as a weather-forecaster in the US Air Force.)
9 Charles Taylor, a prominent defender of the inside approach, says (1985: 116) that social science fails unless it also makes sense of the agents.
10 E.g., Vaitkus 1994: 76, erroneously says, 'a … "realist"… abides by the concerns of natural science'. The doctrine of political neorealism and its most well-known proponent Waltz, has mistakenly been linked to the (unrelated) doctrine of SR (Dessler 1989: 445, Wendt 1987: 351 n. 35;). Waltz is very clearly a scientific anti-realist (Chernoff 1998, 2002: 192).
11 Are *a priori* arguments based on criteria other than consistency always likely to be fruitless? Leibniz's arguments against absolute space and time were certainly not *a posteriori*, based as they were on thought experiments; they were, however, ultimately shown to be correct. They stemmed from his rejection of any 'distinction without a difference' and violated his principle of the identity of indiscernibles, which is ordinarily taken to be a metaphysical principle. See *The Leibniz–Clarke Correspondence* (Leibniz 1956) and Chernoff 1981.
12 All sides agree on simplicity as a desideratum for scientific theories. But why? Instrumentalists have a good answer: it makes theories easier to use and renders them more efficient. But what about scientific realists? How do they answer this question? Older realists argued that God would not make a complex set of laws or forces when He could make a simpler one that did the same things. It was God's intelligence that was the grounding for the economy of laws of nature … and the ontology of the more convoluted theory might be the 'true' one rather than the ontology of the simpler one. Critical realists support the criterion of 'simplicity', but are unusual in placing a lower priority on it. See Patomäki 2002: 148.

13 It was noted above that many social science theorists endorse some but not all of the tenets of natural science in construing social science theory. Non-naturalists reject all of those tenets; those who accept some or most are semi-naturalists. A thoroughgoing naturalist would presumably endorse all of them.

3 Theory, observation and law

1 See the discussion of the position of Little (1993a, 1993b) below in Chapters 4 and 5. Cartwright (1983: 75) follows Michael Scriven (1962) and Sylvan Bromberger (1966) in drawing this distinction between two types of explanatory laws. Hempel and Oppenheim (1965: 267) distinguish 'fundamental' from 'derivative' laws.

2 For Carnap a theoretical law includes both micro-laws and macro-concepts (Carnap 1966: 227–8). On privileged groups, see Olson 1965.

3 LB2 and LB3 assume that theorists in different traditions within a discipline can disagree about why states act as LB2 or LB3 describe, but the IP thesis advanced by Kuhn (1962) denies the presupposition that the theorists agree on a single understanding of concepts like 'public good', hegemonic system', 'democratic state' and 'war'. The IP thesis is explored in Chapter 6.

4 The conventionalist meta-theory espoused in this study is compatible with this caveat, since it has no qualms about the coherence of this causal approach, as the latter deals with observable entities and the terms that refer to them.

5 Dretske (1969: 205 n.1) notes that there are also some differences between 'see' and 'observe'. In general, the analysis would apply *mutatis mutandis* to other senses; 'I observe/see that you have changed your brand of perfume'. Dretske's definition of 'observable' is as follows: x is observable if and only if x can be seen (i) in a non-epistemic way, (ii) in a primary epistemic way or (iii) in a secondary epistemic way (1969: 203).

6 Three major sorts of argument that entail scepticism of the applicability of IR theory stem from: limitations on 'Newtonian conceptions' in the physical sciences, limitations on the parallels between the natural and social sciences, and limitations on what can be known in parallel disciplines in the natural or social sciences. The first set of limitations have been argued in the form of the radical underdetermination principle, the uncertainty principle, IP and CS (all discussed further in the next three chapters). The second sort stem from the HT and reflexivism, and the third sort stem from critical realism, which does not distinguish scientific causal enquiry from social science enquiry but rather subsumes scientific causal enquiry under reflexive studies.

7 Scientific realists and critical realists argue that IR theorists broadly agree on their ontologies, as is shown elsewhere (Chernoff 2002). Although those two doctrines are rejected by this study, they are right on this particular point, which strengthens the claim here about the minimal practical effects of the theory-ladenness of observation thesis.

8 Cartwright distinguishes sharply between non-causal, i.e., mathematical, and causal explanatory laws in science. Both sorts have a proper role but inference to the best explanation is only warranted in the latter case. Some, like Little (1993a), hold that all proper explanation in science is causal, and others, like Russell in *Mysticism and Logic* (1918), hold that none is. Cartwright's position here is intermediate between these two.

9 What is especially interesting about the philosophical framework from which Cartwright presents her argument against the literal truth of theoretical laws is that she believes that theoretical entities exist. Cartwright is not the first to argue that the natural sciences endorse theories that include literal falsehoods. Descartes makes a similar point, e.g., about the epicycles system in astronomy. Descartes argues that false suppositions are sometimes useful for science. See Clatterbaugh 1999: 58, who cites Descartes 1988 v. III: 107.

10 Vasquez endorses the Lakatosian view according to which the superior theory i) has excess empirical content, i.e., can predict novel facts; ii) explains unrefuted content of the older theory; and iii) includes excess content some of which is corroborated (1998: 28). He endorses the criteria of: accuracy, falsifiability, explanatory power, progressive research programme, consistency with other fields of knowledge, parsimony/elegance (Vasquez 1998: 230). In Kuhn's later statement on the subject, he endorses '[a]ccuracy, scope, simplicity, fruitfulness, and the like' (Kuhn 1962: 261).

11 Popper has put much effort into working out problems between his endorsement of SR and the criterion of simplicity – but without fully reconciling the two (see Sober 1988 and above p. 221 note 11).

12 The different definitions of 'motion' do not undermine rational theory choice, since the physics community agreed on the Newton definition because of the agreement on the 'conservation' desideratum; and they do not undercut Dretske's defence of the 'observation–theory' distinction, since his argument does not assert that all observation is theory-neutral, only that some is.

13 Peirce, who was trained as a chemist, insisted also that his scientific methodology conforms to the experimentalist's point of view. The pragmatic theory of truth evidently does not rule out an independent existing reality, since Peirce endorsed both.

4 Natural causation, social action and international politics

1 Pollack (2002) emphasises and amplifies the difficulties of nation-building.

2 See Salmon 1989: 18–20. On the distinction between 'fundamental and derivative laws', see Hempel and Oppenheim 1965: 267.

3 Patomäki (2002: 93 n. 8) makes reference to this sort of context-dependence. Suganami (1996: 130–1), in presenting his argument about necessary conditions, acknowledges that when one seeks the cause of Z's death, the answer 'Z was born' is insufficient and is not properly a cause, even though Z would not have died had Z not been born. 'For example, in response to a question "what caused Z's death?", we would not seriously reply "his birth," or even "his birth among other things."' Suganami adds, 'His birth therefore is taken for granted, requires no mention, and hence does not even count eve as "*a* cause" in the context of *this* question'.

4 A similar point has to do with the perspective of an investigator; a motorway accident may be the result of half a dozen different factors such that had any one been removed, the accident should not have occurred: a partially distracted driver chatting on a mobile phone; 80 per cent worn brake linings; rain on the road surface; the driver travelling at the rate of 75 kilometres per hour in a 60 k.p.h. zone. A psychologist, brake manufacturer, highway engineer and police officer all looking at the accident may each give different answers when each is asked, 'What is the cause of the accident?' (Carnap 1966: 192).

5 In a letter of 16 January 1954, cited by Fine (1996: 88), Einstein says, 'On this account it can never be said with certainty whether the objective world "is causal". Instead one must ask whether a causal theory proves to be better than an acausal one.'

6 Fine (1996: 24) argues that Einstein held that one has to examine the consequence of quantum theory 'especially for macroscopic bodies. [Einstein] argued that the theory was unable to account for (= predict) even the simplest of these phenomena unless we understand the theory never to treat individual systems but only statistical aggregates of such systems.' Fine (1996: 66) also cites Schrödinger's comment (1935: 812) that 'A fuzzy model need not be contradictory … There is a difference between a blurred or out-of-focus picture and a photograph of clouds or fog.'

7 Kyburg makes a similar point about the philosophical treatment of the topic of probability theory. He says that probability theory has been relegated to the sidelines but holds that 'many philosophical problems have been rendered nearly insoluble by the lack of an adequate treatment of probability' (1983: 28).

8 Of course in social sciences like IR there are complaints that some theories simply ignore areas of study that undercut the theoretical claims; this is a practice that should not be permitted. Richard Little points out how state-centric IR theory has ignored studies of the Roman empire, which casts doubt on some central claims (Buzan et al. 1993: 98).

9 Patomäki (2002) argues that fear of error should not deter the development of IR theory. He criticises empiricists who deny causality altogether as inappropriately timid. Yet the fear that causal generalisations will sometimes be mistaken is the sole basis for his rejection of them.

10 Despite the difference in their choice of terms, Little's distinction is clearly based on Cartwright's, especially in view of the bona fide causal status of 'governing regularities' for Little. Little also acknowledges the influence of Cartwright (1983, 1989) as well as that of Salmon (1984).

11 Social regularities are less reliable than those in the natural world according to Little, for a number of reasons: multiple paths of causation; chaos and turbulence in social systems and sensitivity to parameter change; probabilistic causation, in which case the longer the chain, the lower the probability of the dependent state; and the problem of specification of the model from the theory, as the same theory can generate several mutually exclusive models.

12 'At its core, real explanation is always based on causal inferences' (King et al. 1994: 75 n.1).

13 The disagreement may be a result of a different meaning that Wendt and Little attach to the term 'mind-independence'. Wendt seems to apply it to the independence of the mind of an individual investigator, while Little seems to apply it to the independence of human minds generally.

14 Little (1991) also argues that all social explanation is causal. This, conjoined with Cartwright's claim that abductive inference is valid in causal explanation, yields the conclusion for Little that abduction is generally valid in the social sciences. Little (1994: 484–5, 1998: 12–13) shows that functional and collective action explanations are causal.

15 The section on pp. 119–23 below argues that 'causation' in the social context is the core and not derivative sense of the term.

16 It will be argued in the next chapter that the relativisation to a body of evidence does *not* constitute any radical scepticism of the sort attributed to the IP thesis of Kuhn.

17 Social ontologies are fundamentally different from physical ontologies, as the latter are 'integrated' or 'univocal' in the following sense. Molecular biology posits DNA strands and chromosomes, astronomy posits planets, magnetism and gravity, and atomic physics posits neutrinos, positrons and leptons. But the philosopher of science does not have to choose which ontology to accept or 'believe in' in the way that social scientists do. Physical science offers an account in which the physical objects are integrated: DNA is composed of atoms, which are composed of subatomic particles; magnetic forces operate in terrestrial contexts as well as celestial ones and planets are composed of organic and inorganic matter that is composed of DNA, carbon, aluminium, sodium and other atomic constructs.

In IR different sorts of theories are invoked for different sorts of problems: military power is crucial for deterrence theories, economic power and the world capitalist system are crucial for various theories of political economy, norms are important for theories of regimes and laws, and bureaucratic roles and core missions are central to bureaucratic theories of foreign policy. This much is parallel to the natural sciences. However, when it comes to seeing how the ontologies relate to one another, the parallel ends. There is no obvious way of combining these in a 'unified ontology' – and there is no apparent need to do so. They are orthogonal to one another. While causal explanations must be plausible in order for a social science theory or model to reach a threshold of consideration for acceptability (i.e., it is a necessary condition), there is no reason to place as much scientific emphasis on the causal mechanism.

18 On Leibniz's arguments see Leibniz 1956 and Chernoff 1981.

19 Little (1991) makes a good case for a form of methodological individualism and for causal realism. Chapter 6 argues that Little's form of realism is consistent with a notion of 'prediction' grounded in what policy-makers actually do – as opposed to the more grandiose conception Little seems to have in mind.

20 Suganami agrees with Hobbes as well as Kneale (1949) and Scriven (1975) in saying that to explain the occurrence of a given event is to render its occurrence more intelligible than before by solving specific puzzles we have about it. This proposition captures that which the different sorts of causal explanation have in common.

21 The Greeks were the first to attempt to take, 'as we are bound to call it … a *scientific* grip upon reality', which was:

> still imbued with the idea of values stemming from the social sphere. To the extent that this ethico-social sphere, dominated by religious and conservative ideas, must be accepted as a fixed datum, the enquiring mind, the pure striving for knowledge – which in higher and relatively stabilised social conditions flourishes more vigorously alongside the emotional components of consciousness – turns towards the reality perceptible to the senses; and all the more so in that the Greek popular religion presented few obstacles in this direction.
>
> (Kelsen 1973: 167)

22 Kelsen says:

> The cosmos was ruled by a set of laws given by a rational intelligence: nature by god(s) and human affairs by the state and humans. When a genuine dualism developed, the status of these diverged; the laws of nature became necessary and mechanical. But the changes in our understanding of nature from the primitive or pre-scientific understanding (such as that found today in pre-scientific social groups) shows that the scientific notion of causality is not an *a priori* category of thought. It arose, nearly disappeared and reappeared, to be called into question once again with the advent of the quantum theory. The process of increasing scientific objectivity continues with the anthropo-centric or socio-centric view being replaced by an objective order of nature and eventually in astronomy with the geo-centric view disappearing altogether.
>
> (1973: 201)

23 For the probability assignment p, $0<p<1$. Otherwise the statement will not be revisable. An essential feature of a fallibilist theory of knowledge is that the theory be non-monotonic; it must be possible to remove a statement from the evidence corpus after it was included. This would be impossible by most formulae that are candidates for carrying out the relevant calculations, since the assignment of a probability of 1 would make a reduction impossible (Kyburg 1988, 1990b).

5 Prediction, theory and policy-making

1 It is worth noting that, with respect to the discussion of Chapter 2, logical empiricists like Reichenbach, Hempel and Feigl were all scientific realists, at least at some point, while logical positivists like Carnap were not.

2 Patomäki seems to regard social prediction as so implausible that even when he lists the anti-positivist grounds for attacking positivism (2002: 4) he does not, as noted above, bother to mention 'prediction'.
3 The post-positivist distaste for social science prediction is clear, whether implicit, as in many theories, or explicit, as with Charles Taylor (1985), who, as noted in Chapter 1, goes so far as to say that prediction in the human sciences is not only impossible, it is 'radically impossible'.
4 Lachmann (1971) contends that Weber's approach has much older roots and is, in its essence, the classical method of scholarship. He sees Weber as interpreting texts in a way that pre-dates behaviouralism and positivism in the social sciences. The traditional response, when one is puzzled by the meaning of a text, whether it is religious, literary, legal, etc., is to seek to ascertain what the author 'meant by it' (Lachmann 1971: 17–18).
5 Bohman (1993: 110) argues that there are at least five types of interpretive circularity, since theoretical descriptions are always subject to problems of selectivity, perspective, incompleteness, unspecified assumptions and parts-whole circularity.
6 An example is the subjective 'insight' criterion of theory choice presented by Taylor (1985).
7 Suganami appears to endorse predictiveness in IR. He does not discuss it explicitly, but his analysis implies its possibility. He is a strong proponent of 'causation' in the social sciences, a helpful feature for any positive account of 'predictiveness'. And one element in his analysis rests on 'manipulability, in principle', which presupposes predictive consequences of understanding 'causation' (Suganami 1996: 131–4).
8 Scientific realists generally hold that the laws of good scientific theories are approximately true and that the entities referred to therein, whether observable or unobservable, are real. See the discussion in Chapter 2 above and Chernoff 2002.
9 See Chapter 4 note 10. Little contrasts 'governing and phenomenal' regularities rather than 'fundamental and phenomenological' ones, as Cartwright does. He holds, further, that microfoundations are part of any complete causal explanation (Little 1998: 197–8).
10 On the question of the state as an actor, we note that Little denies the existence of super-individuals in the social world (Little 1998: 197–8).
11 While endorsing a strong notion of 'objectivity' in the social sciences, Little himself opposes any notion of the unity of the social sciences: 'Each discipline has its own sophisticated methods of enquiry through which the scientist is able to probe the phenomena of interest' (1993b: 178).
12 Little believes his account is consistent with naturalism (1993a: 196). But he is perhaps more than merely 'consistent' with naturalism because of the narrow, natural science-like standards he invokes for social explanation (causal only) and prediction. The concepts 'explanation' and 'prediction' can be understood in a more inclusive way in the context of the social sciences.
13 Indeed, the impulse to seek regularities in the social world stems from a defensible source, the desire to construct a powerful explanatory and predictive social theory. If further examination shows this project (or portions thereof) impossible, then it (or those parts) must be abandoned. But, given the impressive accomplishments of modern physical theory, one would seem highly justified in seeking to pursue naturalism as far as results of foundational enquiry permits. Moreover, whether an author is a naturalist is not answered simply with 'yes' or 'no'. It is very complex and multidimensional doctrine. There are a variety of variations of naturalism, each having several different elements of the core of the doctrine (see Bhaskar 1998, Hempel 1965, Putnam 1975).
14 Predictions, according to Little, frequently fail also because they are subject to *ceteris paribus* conditions, models are based on simplifying assumptions (which are literally

false), social causal fields are highly complex, and individuals or whole populations are capable of non-rational action (1991: 225–36).

15 Doran uses the term 'forecasting', which he defines as 'a prediction based on knowledge of past behavior' (1999: 12). He does not define 'prediction' but his definition of 'forecasting' comes close to the definition of 'social prediction' as a (probabilistic) statement about the future based on a rational corpus of beliefs, which was proffered in the definition of 'prediction' on p. 8.

16 Despite their inability to generate accurate predictions, dynamical systems models are nevertheless important, according to Doran, because of their capacity for the construction of explanatory models.

17 Doran cites his own statement of this example (1980) and a version that appears in Harvey 1989.

18 Doran regards 'prediction' and 'expectation' as equivalent (1999: 11).

19 It must be remembered that a non-linearity on A's curve may be more a result of new policies of B than of A. Both the slope and the coordinates of the point of the curve are relative to what the rest of the actors in the system are doing.

20 Because Doran uses power-cycle theory, it is appropriate to place an evaluation of his view of 'prediction' in the context of that theory. However, some other as-yet-undiscovered explanatory framework might show the effect of inaccurate prediction by states to play a greater causal role.

21 This example is simplified by using a single desired outcome, T, in all three propositions, which will then have the same utility in all cases. In more complex cases, the expected utility calculations will involve variability in both the probability and utility values, which are then combined to yield an 'expected utility' value for the course of action. Likewise, in this simplified case, we may assume that the confidence levels of all three probability assignments are equal.

22 While many have made use of the structure of biological theory, especially evolutionary biology, as a model for IR theory, some have sought to use evolutionary biology as a foundation for the premises of IR theory. Thayer (2001) uses evolution as a grounding for classical-realist premises about the drive to dominate.

23 Keohane clearly counts military alliances in general as institutions. He says, 'a variety of international institutions, including most obviously military alliances, are designed as a means for prevailing in military and political conflict' (Keohane 1989: 159). Keohane says, ' "*institution*" may refer to a *general pattern* or *categorization* of activity or to a *particular* human constructed arrangement, formally or informally organized' (Keohane 1989: 162, italics in original). Keohane (1989: 163) points out that 'Douglass North (1987: 6) defines institution as "rules, enforcement characteristics of rules, and norms of behavior that structure repeated human interaction."' I thank Nate Webb for the citations.

24 Supporters of the IP thesis of Kuhn (1962) would deny that there is commensurability of the terms of neorealism and neoliberalism. But as one can see, the definitions are often close enough that the key disagreements between different theorists are phrased in such a way that both are talking about the same objects classified in the same ways by the terms they use.

25 The insistence on being able to make future policy while denying prediction seems incongruous. This is a problem for any account that acknowledges the imperative of prediction for policy-making but rejects the predictiveness of IR theory. The first sentence of the quotation suggests the quotation from Ibsen at the head of this chapter.

26 In a technical sense, there may be 'necessary truths' as part of IR prediction for those who endorse the 'partial entailment' interpretation of probability statements. For example, the prediction 'there is a 30 per cent chance that treaty x will be signed in the next decade, given the available evidence' would constitute a necessary truth, since it expresses a logical relationship between a body of evidence and the predic-

tion. But this does not fall victim to the sort of criticism Bernstein et al. develop, since such a prediction is clearly not deterministic.

27 Popper (1971b: 322 n. 13) says:

> There can be *sociological* laws, and even sociological laws pertaining to the problem of progress; for example, the hypothesis that, wherever the freedom of thought, and the communication of thought, is effectively protected by legal institutions and institutions ensuring the publicity of discussion, there will be scientific progress … But there are reasons for holding the view that we should do better not to speak of *historical* laws at all.

Popper distinguishes 'generalising' social sciences like sociology from the historical sciences and holds that both include causal explanations (1971b: 264).

28 This seems to conform to practice in IR. For example, in the most vigorously contested area of empirical enquiry in IR at this time, DP studies, the most recent work by Russett and Oneal concludes as follows:

> [s]upport for the Kantian peace, especially the benefits of democracy and economic interdependence , is extremely robust, as we have tried to show in this book. No variable is significant in every possible test, but it the weight of the evidence is considered, there is little doubt that the liberals were right: democracy, economic interdependence, and cooperation in international organizations reduce the incidence of war.
>
> (Russett and Oneal 2001: 313)

6 Explaining agreement and disagreement in the natural sciences and social sciences

1 Constructivists classify both neorealists and liberals as 'rationalists' and distinguish themselves from both traditions by emphasising the constitutive over the causal relationships in international theory and by arguing that international concepts are constructed rather than externally given (see Wendt 1992, 1995). Constructivists still need to explain why dissensus remains on fundamental questions in IR. Many scholars in IR who argue in the constructivist vein and those sympathetic to critical theory have re-injected a moral dimension into the study of IR, holding that the purpose of the field of IR and other social sciences, as well as the natural sciences, is 'human emancipation'; see Wendt (1987), Williams and Krause (1997), and Wyn-Jones (1999).

2 Approach-to-consensus does not mean a steady progression throughout the history of physics but rather movement toward consensus after the previous consensus is disrupted by introduction of a new paradigm or disciplinary matrix. Within the philosophy of the physical sciences there is debate about 'rationality' and 'progress'. Philosophers and historians of science like Kuhn (1962) do not see physics as a rational progression from one theoretical paradigm to another, superior paradigm. But Kuhn does acknowledge that physical science moves towards consensus when certain new paradigms are introduced. In contrast, Lakatos sets out the goal of producing 'a rational explanation of the growth of objective knowledge' (1970: 102).

3 Naturalist-leaning authors like Vasquez see consensus in IR as a serious possibility. He says (1998: 29) 'As more research is conducted and more evaluations of it are made, a trend may become clear and the disagreements will probably subside.' But even some reflectivists seem inclined to hope for approach-to-consensus in IR. Critical realism is one example. Patomäki says, 'The result of dialectical exchange does not consist of

purely negative or contradictory countermoves; it advances the discussion and shifts the issue onto more sophisticated ground' (2002: 71). The goal for Patomäki is 'a way out of the increasingly outmoded "great debates" ' (2002: 73).

4 Majer (1995), in supporting Hilbert's account, attacks this view and argues that Einstein overlooks geometry's unique position of standing between logic and experience. See Appendix to this chapter.

5 As the Appendix shows, Carnap's aim was the reconciliation of Kant's commitment to space as a form of intuition with both relativistic physics and the retention of Euclidean geometry for the world of experience. Carnap is thus concerned with a particular theory, and that theory cannot be chosen in a way entirely independent of conventional – non-theoretical – decisions. How do we choose one among the many Riemannian manifolds that we are presented with as a synthetic *a priori* form of intuition? Carnap acknowledges that this can only be accomplished through a convention: either a 'measure-stipulation' namely, that bodies used as a measuring standard remain rigid when transported, or, alternatively an equally conventional choice directly of one of the geometric axiom systems.

6 The measure-stipulation only adjudicates between a limited number of theories. Most theories will be shown to be inferior to the best theory by means of other criteria. In physics, many theories are consistent with the stipulation that a measuring rod remains of constant length when it is moved in space. Most theories with laws that vary from Einstein's will fail to accord with various astronomical observations, which accord with Einstein's. On those grounds innumerable other theories may be rejected. The measure-stipulation is a crucial element of physical theory but this conventional move is not the only, or even the most important, element in theory choice. And this conventional move is one that is taken on the basis of rational criteria, as Poincaré and Duhem argue.

7 Friedman 1992: 248. This is the only way available for Carnap to move from the Euclidean manifold of intuitive space to the non-Euclidean manifold of relativistic physical space. There is a wide choice of non-Euclidean manifolds that can be brought into line with observed phenomena. The choice of a single manifold comes about by stipulation. See Friedman 1992: 251.

8 Friedman (1983: 304) argues that one may only conclude from Grünbaum's observations that the metric is a primitive or undefined quantity and that another premise about topology is needed to prove the conclusion about metrical amorphousness.

9 Margaret Masterman's (1974) oft-quoted paper counts twenty-one different senses of the term in Kuhn (1962).

10 The recent critique of incommensurability by Wight (1996) also makes a number of powerful points.

11 This is, for example, one of the reasons Kuhn (1962) does not extend his account of the history of natural sciences, where genuine paradigms are accepted throughout a discipline, to the social sciences, even though social scientists have almost unanimously ignored this qualification.

12 This is one variation of the many statements by Chisholm of his epistemic principles stated in *Perceiving* (1957: e.g., 151), evolving through the three editions of *Theory of Knowledge* (1966, 1977 and 1989), and in his recent work (e.g., 1996).

13 Recent liberals who do examine security most often focus on the differences between democracies and non-democracies, not on power relations. See pp. 189–92.

14 The term 'DP hypotheses' and 'proponents' is used generically for the various propositions asserting, and for those who hold, that democratic, republican or liberal regime type explains observed patterns of peace and war.

15 A number of authors argue in favour of scalar tests. Elkins (2000: 299) shows that they are more valid in that 'measures of democracy which provide for gradations best fit the behaviour that theoretical work on democracy would predict'.

16 Russett's (1995: 168–9) response to Spiro is not as progressive, essentially charging that Spiro carries out the same sort of altering of correlates of war definitions as that Spiro accuses Russett of doing.

17 Gartzke (1998: 6) says:

> [o]bservations of the democratic peace are not unlike studying incidents of seasickness in Central Asia. There is nothing to report. Still, we cannot then assume that Uzbek culture makes them hearty seafaring folk or that Tadzhik bureaucrats introduce a mysterious regime that makes local villagers immune to the effects of vertigo.

18 Beck et al. (1998: 1279) say, '[c]ounting the latter years of a multi-year disputes as new disputes, and failing to correct for dependence between these disputes, is what leads to the Oneal and Russett finding that trade lowers the probability of the onset of a dispute'.

19 Green et al.'s (2001) critique uses data that go from 1951 until 1992, while the data that Oneal and Russett analyse begin in 1886 and go up to 1992 using the same data as Green et al. for the post-war period. The two papers also achieve different results because, when they assess trade volume, they differ on 'a seemingly minor methodological decision: how to treat zero levels of trade when taking a logarithm' (Oneal and Russett 2001: 469.)

20 Weart (1994: 310) argues for three years. Huntington (1991) argues for a definition in terms of personnel-turnover.

21 Elsewhere (Chernoff 2004) it is argued that the democratic peace literature provides an example of scientific progress in IR, regardless of which major account of 'science' and 'progress' one adopts in the philosophy of science.

(22 Chan holds that, '[o]ne clear-cut case of … belligerence would be sufficient to disconfirm the democratic peace proposition' (Chan 1997: 71; see also Rummel 1983: 29), which is inconsistent with the CC framework. Much closer is the work of Oneal and Russett (2001: 216), who hold that probabilistic laws are not falsified by a single case.

23 Rosenkranz (1981) offers an interesting argument defending the synthetic *a priori* nature of one form of space.

7 Conclusions

1 See also Harré, 1986: 89, as cited by Patomäki 2002: 148. This study agrees with the 'moral principle' articulated there, if not with all the reasons for accepting it.

2 This is the heading of section Ten of Duhem's *The Aim and Structure of Physical Theory* (1954).

3 Still, Duhem's conventionalist holism requires that an accumulation of such instances shows that the theory is inferior to its rivals.

4 Falsifiability is closely related to external consistency, but the former is logical and the latter is not, since the latter is about the nature of the theory and its relationship to *possible* observations.

5 I thank Dan Nexon for bringing to my attention the consequences of this argument.

6 See also Linklater's (1990) related account of pluralism.

Bibliography

Adorno, Theodor W. (2000) *The Adorno Reader*. Trans. and ed. Brian O'Connor. Oxford: Blackwell.

Agarwal, G.S. (1974) *Quantum-Statistical Theories of Spontaneous Emission and their Relation to Other Approaches*. Berlin: Springer-Verlag.

Alexander, Peter (1963) *Sensationalism and Scientific Explanation*. London: Routledge & Kegan Paul.

Alston, William P. (1990) *Epistemic Justification: essays in the theory of knowledge*. Ithaca, NY: Cornell University Press.

Aristotle (1956) *Physics*. Oxonii: E. Typographeo Clarendoniano.

—— (1984) *The Complete Works of Aristotle*. Ed. Jonathan Barnes. Princeton, NJ: Princeton University Press.

Arquilla, John (1992) *Dubious Battles: aggression, defeat and the international system*. Washington, DC: Crane Russak.

Arreguín-Toft, Ivan (2001) 'How the Weak Win Wars: a theory of asymmetric conflict' *International Security* 26: 93–128.

Ashley, Richard C. (1986) 'The Poverty of Neorealism'. In *Neorealism and Its Critics*. Ed. Robert O. Keohane. New York: Columbia University Press, pp. 255–300.

Audi, Robert (1993) *The Structure of Justification*. Cambridge: Cambridge University Press.

Babst, Dean (1964) 'Elective Governments: a force for peace' *Wisconsin Sociologist* 3: 9–14.

—— (1972) 'A Force for Peace' *Industrial Research* April: 55–8.

Barkawi, Tarak and Mark Laffey (2001) 'Introduction: the international relations of democracy, liberalism, and war'. In *Democracy, Liberalism and War: rethinking the democratic peace debate*. Eds. Tarak Barkawi and Mark Laffey. Boulder, Col.: Lynne Rienner, pp. 1–24.

Barton, Frederick and Bathsheba Crocker (2003) 'Winning the Peace in Iraq' *The Washington Quarterly* 26: 7–22.

Bayes, Thomas (1763) 'An Essay Toward Solving a Problem in the Doctrine of Chances' *Philosophical Transactions of the Royal Society* 53: 370–418.

Beck, Nathaniel and Jonathan N. Katz (2001) 'Throwing the Baby Out With the Bathwater: a comment on Green, Kim, and Yoon' *International Organization* 55: 486–500.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker (1998) 'Beyond Ordinary Logic: taking time seriously in binary-time-series-cross-section models' *American Journal of Political Science* (42) 4: 1260–88.

Bernstein, Steven, Richard Ned Lebow, Janice Gross Stein and Steven Weber (2000) 'God Gave Physics the Easy Problems: adapting social science to an unpredictable world' *European Journal of International Relations* 6: 43–76.

Bhaskar, Roy (1975) *A Realist Theory of Science*. London: Verso.
—— (1978) *The Possibility of Naturalism*. London: Harvester Wheatsheaf.
—— (1986) *Scientific Realism and Human Emancipation*. London: Verso.
—— (1998) *The Possibility of Naturalism*, 3rd edn. London: Routledge.
Bird, Graham (1995) 'Carnap and Quine on Internal and External Questions' *Erkenntnis* 42: 41–64.
Bohman, James (1993) *The New Philosophy of Social Science: problems of indeterminacy*. Cambridge, Mass.: MIT Press.
Bourdieu, Pierre (1977) *Outline of a Theory of Practice*. Trans. Richard Nice. Cambridge: Cambridge University Press.
Boyd, Richard N. (1973) 'Realism: underdetermination and a causal theory of evidence' *Noûs* 7: 1–12.
Bradley, Richard. (1999) 'Review of *Explorations in Economic Methodology*' *British Journal of Philosophy of Science* 50: 316–18.
Bromberger, Sylvan (1966) 'Why Questions'. In *Mind and Cosmos*. Ed. Robert G. Colodny. Pittsburgh, Pa.: University of Pittsburgh Press, pp. 86–111.
Brooks, Rise (2002) 'Liberalization and Militancy in the Arab World' *Orbis* 46: 611–21.
Brown, Courtney (1994) 'Politics and the Environment: nonlinear instabilities dominate' *American Political Science Review* 88: 292–303.
Bueno de Mesquita, Bruce (2002) *Predicting Politics*. Columbus, Ohio: Ohio State University Press.
Bueno de Mesquita, Bruce and David Lalman (1994) *War and Reason*. New Haven, Conn.: Yale University Press.
Bull, Hedley (1977) *The Anarchical Society: a study of order in world politics*. New York: Columbia University Press.
Butterfield, Herbert (1966) 'The Balance of Power'. In *Diplomatic Investigations: essays in the theory of international politics*. Eds. H. Butterfield and M. Wight. London: Allen & Unwin, pp. 132–48.
Buzan, Barry, Charles Jones and Richard Little (1993) *The Logic of Anarchy: neorealism to structural realism*. New York: Columbia University Press.
Byman, Daniel (2003) 'Constructing a Democratic Iraq: challenges and opportunities' *International Security* 28: 47–78.
Capelle, Wilhelm (1938) *Die Vorsokratiker*, 2nd edn. No. 25. Stuttgart: A. Kröner
Carnap, Rudolf (1950) *The Logical Foundations of Probability*. Chicago, Ill.: University of Chicago Press.
—— (1956) 'Empiricism, Semantics and Ontology'. In *Meaning and Necessity: a study in semantics and modal logic*, 2nd edn. Chicago, Ill.: University of Chicago Press, pp. 205–21.
—— (1966) *An Introduction to the Philosophy of Science*. Ed. Martin Gardner. New York: Dover.
—— (1967) *The Logical Structure of the World: pseudoproblems in philosophy*. Trans. Rolf A. George. Berkeley, Calif.: University of California Press.
Carr, Edward Hallett (1964) *The Twenty Years' Crisis, 1919–1939: an introduction to the study of international relations*. New York: Harper.
Cartwright, Nancy (1983) *How the Laws of Physics Lie*. Oxford: Clarendon Press.
—— (1989) *Nature's Capacities and Their Measurement*. Oxford: Clarendon Press.
Cederman, Lars-Erik (1996) 'Rerunning History: counterfactual simulation in world politics'. In *Counterfactual Thought Experiments in World Politics: logical, methodological, and psychological perspectives*. Eds. Philip E. Tetlock and Aaron Belkin. Princeton, NJ: Princeton University Press, pp. 247–67.

Chan, Steve (1984) 'Mirror, Mirror on the Wall… Are the Freer Countries More Pacific?' *Journal of Conflict Resolution* 28: 617–48.

—— (1997) 'In Search of Democratic Peace: problems and promise' *Mershon International Studies Review* 41: 59–91.

Chernoff, Fred (1981) 'Leibniz's Principle of the Identity of Indiscernibles' *Philosophical Quarterly* 31: 126–38

—— (1995) *After Bipolarity: theories of cooperation, the vanishing threat and the future of the Atlantic alliance*. Ann Arbor, Mich.: University of Michigan Press.

—— (1998) 'Scientific Realism and Theories of International Relations'. Paper presented at International Studies Association, Minneapolis, MN: (March).

—— (2002) 'Scientific Realism as a Meta-Theory of International Relations' *International Studies Quarterly* 46: 189–207.

—— (2004) 'The Study of Democratic Peace and Progress in International Relations' *International Studies Review* 6: 49–78.

Chisholm, Roderick M. (1957) *Perceiving: a philosophical study*. Ithaca, NY: Cornell University Press.

—— (1966, 1977, 1989) *Theory of Knowledge*, 1st, 2nd and 3rd edn. Englewood Cliffs, NJ: Prentice Hall.

—— (1996) *A Realistic Theory of Categories: an essay on ontology*. New York: Cambridge University Press.

—— (1997) 'My Philosophical Development', and 'Replies to Critics'. In *The Philosophy of Roderick Chisholm*. Ed. Lewis Edwin Hahn. Chicago and La Salle, Ill.: Open Court, pp. 3–41 and passim.

Clatterbaugh, Kenneth C. (1999) *The Causation Debate in Modern Philosophy, 1637–1739*. New York: Routledge.

Clendinnen, F. John (1979) 'Inference, Practice and Theory'. In *Hans Reichenbach, Logical Empiricist*. Ed. Wesley H. Salmon. Dordrecht: D. Reidel Publishing Co., pp. 85–128.

Collingwood, R.G. (1946) *The Idea of History*. Oxford: Oxford University Press.

Comte, Auguste (1970) *Introduction to Positive Philosophy*. Trans. and ed. Frederick Ferré. Indianapolis, Ind.: Bobbs-Merrill.

Cox, Robert W. (1987) *Power, Production, and World Order: social forces in the making of world history.* New York: Columbia University Press.

Cumings, Bruce (2001) 'Democracy and Peace: what's not to love?'. In *Democracy, Liberalism, and War: rethinking the democratic peace debate*. Eds. T. Barkawi and M. Laffey. Boulder, Col.: Lynne Rienner.

Davidson, Donald (1984) *Essays into Truth and Interpretation*. Oxford: Oxford University Press.

Derrida, Jacques (2002) *Without Alibis: selections*. Trans. Peggy Kamuf. Stanford, Calif: Stanford University Press.

Descartes, René (1988) *Descartes: selected philosophical writings*. Trans. and eds. John Cottingham, Robert Stoothoff and Dugald Murdoch. Cambridge: Cambridge University Press.

Dessler, David (1989) 'What's at Stake in the Agent–Structure Debate' *International Organization* 43: 441–73.

Deutsch, Karl et al. (1957) *Political Community and the North Atlantic Area*. Princeton, NJ: Princeton University Press.

Diehl, Paul (1983) 'Arms Races and Escalation: a closer look' *Journal of Political Research* 20: 205–12.

DiFinetti, Bruno (1964) *Foresight: its laws, its subjective sources*. Trans. Henry E. Kyburg, Jnr. In *Studies in Subjective Probability*. Ed. H.E. Kyburg, Jnr. and H. Smokler. New York: Wiley, pp. 97–158.

Dilthey, Wilhelm (1996) *Hermeneutics and the Study of History*. Eds. Rudolf A. Makkreel and Frithof Rodi. Princeton, NJ: Princeton University Press.

Dixon, William J. (1993) 'Democracy and the Management of International Conflict' *Journal of Conflict Resolution* 37: 42–68.

Dobbins, James, John G. McGinn, Keith Crane, Seth G. Jones, Rollie Lal, Andrew Rathmell, Rachel Swanger, and Anga Timilsina (2003). *America's Role in Nation Building: from Germany to Iraq*. Santa Monica, Calif.: The Rand Corporaton.

Donagan, Alan (1964) 'Historical Explanation: the Popper-Hempel theory reconsidered' *History and Theory* 4: 3–26.

Doran, Charles F. (1980) 'Mechanisms and Turning Points: perspectives on the analysis of the transformation of the international system' *International Political Science Review* 1: 35–61.

—— (1989) 'Power Cycle Theory of Systems Structures and Stability: commonalities and complementarities'. In *Handbook of War Studies*. Ed. Manus I. Midlarsky. Boston, Mass.: Unwin Hyman, pp. 83–111.

—— (1991) *Systems in Crisis*. Cambridge: Cambridge University Press.

—— (1999) 'Why Forecasts Fail: the limits and potential of forecasting in international relations and economics' *International Studies Review* 1: 11–41.

Doyle, Michael (1983a) 'Kant, Liberal Legacies, and Foreign Affairs: part 1' *Philosophy and Public Affairs* 12: 205–35.

—— (1983b) 'Kant, Liberal Legacies, and Foreign Affairs: part 2' *Philosophy and Public Affairs* 12: 323–53.

—— (1986) 'Liberalism and World Politics' *American Political Science Review* 80: 1151–61.

Dretske, Fred I. (1969) *Seeing and Knowing*. Chicago, Ill.: University of Chicago Press.

—— (1979) 'Chisholm on Perceptual Knowledge' *Grazer-Philosophische-Studien* 7/8: 253–69.

Dryzek, John (1996) *Democracy in Capitalist Times*. Oxford: Oxford University Press.

Duhem, Pierre (1954) *The Aim and Structure of Physical Theory*. Trans. Philip P. Wiener. Princeton, NJ: Princeton University Press.

—— (1969) *To Save the Phenomena*. Trans. Edmund Dohland and Chaninah Maschler. Chicago: University of Chicago Press.

Durkheim, Émile (1972) *Selected Writings*. Edited and Trans. Anthony Giddens. Cambridge: Cambridge University Press.

Einstein, Albert (1916) 'Die Grundlage der Allgemeinen Relativitätstheorie' *Annalen der Physik* Band 1949 (7): 769–822.

—— (1949) 'Autobiographical Notes'. In *Albert Einstein Philosopher Scientist*. Ed. Paul Arthur Schilpp. Evanston, Ill.: Library of Living Philosophers, pp. 1–95.

Elkins, Zachary (2000) 'Gradations of Democracy? Empirical Tests of Alternative Conceptualizations' *American Journal of Political Science* 44: 293–300.

Elman, Colin and Miriam Fendius Elman (1997) 'Lakatos and Neorealism: a reply to Vasquez' *American Political Science Review* 91: 923–6.

—— (2002) 'How Not to be Lakatos Intolerant: appraising progress in IR research' *International Studies Quarterly* 46: 231–62.

Elshtain, Jean Bethke (2003) *Just War Against Terror: the burden of American power in a violent world*. New York: Basic Books.

Elster, Jon (1989) *Solomonic Judgments*. Cambridge: Cambridge University Press.

Enloe, Cynthia (2000) *Bananas, Beaches and Bases: making feminist sense of international politics*. Berkeley and Los Angeles, Calif.: University of California Press.

Farber, Henry and Joanne Gowa (1997) 'Common Interests or Common Politics?' *Journal of Politics* 57: 393–417.

Feldman, Noah (2003) *After Jihad: America and the struggle for Islamic democracy*. New York: Farrar, Straus & Giroux.

Feyerabend, Paul K. (1978) *Against Method*. London: Verso.

Fine, Arthur (1996) *The Shaky Game: Einstein, realism and the quantum theory*. Chicago, Ill.: University of Chicago Press.

Fodor, Jerry (1984) 'Observation Reconsidered' *Philosophy of Science* 51: 23–43.

Føllesdal, Dagfinn (1979) 'Hermeneutics and the Hypothetico-Deductive Method' *Dialectica* 33: 319–36.

Freeman, Kathleen (1948) *Ancilla to the Pre-Socratic Philosophers*. Cambridge, Mass.: Harvard University Press.

Friedman, Michael (1983) *Foundations of Space-Time Theories: physics and the philosophy of science*. Princeton, NJ: Princeton University Press.

—— (1992) *Kant and the Exact Sciences*. Cambridge, Mass.: Harvard University Press.

Fukuyama, Francis (1992) *The End of History and the Last Man*. New York: Free Press.

Gadamer, Hans-Georg (1976) *Philosophical Hermeneutics*. Trans. David E. Linge. Berkeley: University of California,

Garnham, David (1986) 'War Proneness, War Weariness, and Regime Type, 1816–1980' *Journal of Peace Research* 23: 279–89.

Gartzke, Erik (1998) 'Kant We All Just Get Along? Opportunity, willingness and the origins of the democratic peace' *American Journal of Political Science* 42: 1–27.

Geertz, Clifford (1973) *Interpretation of Cultures: selected essays*. New York: Basic Books.

Giddens, Anthony (1974) *Positivism and Sociology*. London: Heinemann.

Gillies, Donald (1993) *Philosophy of Science in the Twentieth Century: four central themes*. Oxford: Blackwell Publishers.

Gilpin, Robert (1981) *War and Change in World Politics*. New York: Cambridge University Press.

—— (1989) 'The Theory of Hegemonic War'. In *The Origins and Prevention of Major Wars*. Ed. Robert I. Rotberg and Theodore K. Rabb. Cambridge: Cambridge University Press, pp. 15–37.

Gleditsch, Nils Petter (1995) 'Geography, Democracy, and Peace' *International Interactions* 20: 297–323.

Gleditsch, Kristian and Michael D. Ward (1997) 'War and Peace in Space and Time: the role of democratization' *International Studies Quarterly* 44: 1–29.

Gödel, Kurt (1931) 'Überformal unentscheidbare Sätze der Principia Mathematica und wewandter Systeme I' *Monatshefte für Mathematik und Physik* 38: 173–89.

Goldman, Alvin I. (1999) *Knowledge in a Social World*. Oxford: Clarendon Press.

Gowa, Joanne (1999) *Ballots and Bullets: the elusive democratic peace*. Princeton, NJ: Princeton University Press.

Gramsci, Antonio (1994) *Pre-Prison Writings*. Ed. Richard Bellamy. Trans. Victoria Cox. Cambridge: Cambridge University Press.

Green, Donald, Soo Yeon Kim, and David Yoon (2001) 'Dirty Pool' *International Organizations* 55: 441–68.

Greenwood, John D. (1990) 'Two Dogmas of Neoempiricism: the "theory-informity" of observation and the Quine–Duhem thesis' *Philosophy of Science* 57: 553–74.

Grünbaum, Adolf (1968) *Geometry and Chronometry in Philosophical Perspective*. Minneapolis, Minn.: University of Minnesota Press.

Haas, Michael (1995) 'When Democracies Fight One Another, Just What is the Punishment for Disobeying the Law?' Paper presented at the annual meeting of the American Political Science Association, Chicago.

Habermas, Jürgen (1971) *Knowledge and Human Interests*. Trans. Jeremy J. Shapiro. Boston, Mass.: Beacon Press.

—— (1973) *Theory and Practice*. Trans. John Verteil. Boston, Mass.: Beacon Press.

—— (1984) *Theory and Practice*. Trans. Thomas McCarthy. Boston, Mass.: Beacon Press.

—— (1987) *Theory of Communicative Action*. Boston, Mass.: Beacon Press.

Hanson, Norwood Russell (1958) *Patterns of Discovery*. London: Cambridge University Press.

Harré, Rom (1986) *Varieties of Scientific Realism*. Oxford: Blackwell.

Harvey, Andrew C. (1989) *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.

Havas, Peter (1967) 'Foundation Problems in General Relativity'. In *Delaware Seminar in the Foundations of Physics*. Ed. Mario Bunge. New York: Springer-Verlag, pp. 140–1.

Hayek, Friedrich A. von (1973–8) *Law, Legislation, Liberty: a new statement of the liberal principles of justice and the political economy*. Chicago, Ill.: University of Chicago Press.

Heidegger, Martin (1965) *German Existentialism*. Trans. Dagobert D. Runes. New York: Wisdom Library.

—— (1969) *The Essence of Reasons*. Trans. Terrence Malick. Evanston, Ill.: Northwestern University Press.

Heller, Mark (2003) 'Prospects for Creating a Regional Security Structure in the Middle East' *Journal of Strategic Studies* 26 (3): 125–36.

Hempel, Carl G. (1962) 'Deductive-Nomological versus Statistical Explanation'. In *Minnesota Studies in the Philosophy of Science*. Vol. III. Ed. Herbert Feigl. Minneapolis, Minn.: University of Minnesota, pp. 98–169.

—— (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.

——(1988) 'Limits of a Deductivist Construal of the Function of Scientific Theories'. *In Science In Reflection: The Israeli Colloquium*. Vol 3. Ed. E. Ullman-Margalit. Dordrecht: Kluwer, pp. 1–15.

Hempel, Carl G. and Paul Oppenheim (1965) 'Studies in the Logic of Explanation'. In *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Ed. by Carl G. Hempel. New York: Free Press, pp. 245–91.

Hesse, Mary (1980) *Revolutions and Reconstructions in the Philosophy of Science*. Bloomington, Ind.: Indiana University Press.

Hewitt, J. Joseph and Jonathan Wilkenfeld (1996) 'Democracies in International Crisis' *International Interactions* 22: 123–42.

Hitchcock, Christopher Read (1992) 'Causal Explanation and Scientific Realism' *Erkenntnis* 37: 179–96.

Hobson, John A. (1926) *Free Thought in the Social Sciences*. New York: Macmillan.

Hollis, Martin and Steve Smith (1991) *Explaining and Understanding International Relations*. Oxford: Clarendon Press.

Hollis, Rosemary (2003) 'Getting Out of the Iraq Trap' *International Affairs* 79: 23–35.

Hookway, Christopher (1985) *Peirce*. London: Routledge & Kegan Paul.

Hopf, Ted (forthcoming) 'The Limits of Interpreting Evidence'. In *Social Knowledge*. Eds. Richard Ned Lebow and Mark Irving Lichbach.

Horkheimer, Max (1993) *Between Philosophy and Social Science: selected early writings*. Trans. G. Frederick Hunter, Matthew S. Kramer, and John Torpey. Cambridge, Mass.: MIT Press.

Hudson, Valerie M. (2001) 'Manifest Agency, Foreign Policy Analysis, and IR Theory: the other agent–structure debate and its (highest) stakes'. Paper presented at the 42nd annual conference of the International Studies Association, Chicago, Ill., 20–4 February 2001.

Hume, David (1965) *A Treatise of Human Understanding*. Ed. L.A. Selby-Bigge. Oxford: Clarendon.

Hunt, Shelby D. (1994) 'A Realist Theory of Empirical Testing: resolving the theory-ladenness/objectivity debate' *Philosophy of the Social Sciences* 24: 133–58.

Huntington, Samuel P. (1991) *The Third Wave: democratization in the late twentieth century*. Norman, Okla.: University of Oklahoma Press.

Husserl, Edmund (1913) *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie*. Halle: Max Niemeyer.

Jackson, Patrick Thaddeus and Daniel H. Nexon (1998) 'Relations Before the States: An Inquiry into Sovereignty and International Relations', Paper presented at International Studies Association, Minneapolis, MN: (March),

Jaeger, Werner (1939) *Paideia: the ideals of Greek culture*. Trans. Gilbert Highet. Oxford: Basil Blackwell.

Jaggers, Keith and Ted Robert Gurr (1995) 'Tracking Democracy's Third Wave with the Polity III Data' *Journal of Peace Research* 32: 469–82.

James, Patrick (2002) *International Relations and Scientific Progress: structural realism reconsidered*. Columbus, Ohio: Ohio State University Press.

Jeffrey, Richard C. (1992) *Probabilities and the Art of Judgement*. Cambridge and New York: Cambridge University Press.

Jenkins, Rhys (1987) *Transnational Corporations and Uneven Development: the internationalization and the third world*. London: Routledge.

Jervis, Robert (1976) *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.

—— (1997) *System Effects: complexity in political and social life*. Princeton, NJ: Princeton University Press.

Jørgensen, Knud Erik (2001) 'Four Levels and a Discipline'. In *Constructing International Relations: the next generation*. Eds. K.M. Fierke and K.E. Jørgensen. Armonk: M.E. Sharpe, pp. 36–53.

Kant, Immanuel (1950) *Critique of Pure Reason*. Trans. Norman Kemp Smith. New York: Humanities Press.

Kaplan, Morton A. (1957) *System and Process in International Politics*. New York: Wiley.

Keat, Russell (1981) *The Politics of Social Theory, Habermas, Freud, and the Critique of Positivism*. Oxford: Basil Blackwell.

Kegley, Charles W. and Margaret Hermann (1995) 'Military Intervention and the Democratic Peace' *International Interactions* 21: 1–21.

—— (1997) 'Putting Military Intervention into the Democratic Peace' *Comparative Political Studies* 30: 78–107.

Kelsen, Hans (1973) 'The Emergence of the Causal Law from the Principle of Retribution'. In *Essays in Legal and Moral Philosophy*. Trans. Peter Heath and ed. Hans Kelsen. Dordrecht and Boston, Mass.: D. Reidel, pp. 165–215.

Keohane, Robert O. (1984) *After Hegemony*. Princeton, NJ: Princeton University Press.

—— (1989) *International Institutions and State Power: essays in international relations theory*. Boulder, Col.: Westview Press.

—— (2000) 'Ideas Part-Way Down' *Review of International Studies* 26: 125–30.

King, Gary, Robert O. Keohane and Sidney Verba (1994). *Designing Social Inquiry: scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.

Kleene, Steven C. (1952) *Introduction to Metamathematics*. Princeton, NJ: Van Nostrand.

Kneale, William (1949) *Probability and Induction*. Oxford: Clarendon.

Knorr-Cetina, Karin et al. (1993) 'The Realism–Constructivism Debate'. In *Taking the Naturalistic Turn: or how real philosophy of science is done*. Ed. Werner Callebaut. Chicago, Ill.: University of Chicago Press, pp. 169–89.

Kratochwil, Freidrich (2001) 'Constructivism as an Approach to Interdisciplinary Study'. In *Constructing International Relations: The Next Generation*. Eds. Karin M. Fierke and Knud Erik Jørgensen. Armonk: M.E. Sharpe, pp. 13–35.

Kripke, Saul A. (1963) 'Semantical Analysis of Modal Logic I: Normal Modal Propositional Calculi', Zeitschrift für Mathematische Logik und Grundlagen der Mathematik 9: 67–96.

Kugler, Jacek and Douglas Lemke (1996) *Parity and War: evaluations and extensions of the war ledger*. Ann Arbor, Mich.: University of Michigan Press.

Kugler, Jacek and A.F.K. Organski (1989) 'The Power Transition: a retrospective'. In *Handbook of War Studies*. Ed. Manus I. Midlarsky. Boston, Mass.: Unwin Hyman, pp. 171–94.

Kuhn, Thomas S. (1962) *The Structure of Scientific Revolutions*. Chicago, Ill.: University of Chicago Press.

Kukla, André (1996) 'Antirealist Explanations of the Success of Science' *Philosophy of Science* 63 (Supplement): S298–303.

Kyburg, Henry E. Jnr. (1961) *Probability and the Logic of Rational Belief*. Middletown, Ohio: Wesleyan University Press.

—— (1977) 'A Defense of Conventionalism' *Noûs* 11 : 75–95.

—— (1983) *Epistemology and Inference*. Minneapolis, Minn.: University of Minnesota Press.

—— (1988) 'Full Belief' *Theory and Decision* 25: 137–62.

—— (1990a) *Science and Reason*. New York: Oxford University Press.

—— (1990b) 'Theories as Mere Conventions'. In *Minnesota Studies in the Philosophy of Science*, 14, *Scientific Theories*. Ed. C. Wade Savage. Minneapolis, Minn.: University of Minnesota Press, pp. 158–74.

Lachmann, L.M. (1971) *The Legacy of Max Weber: three essays*. Berkeley, Calif.: Glendessary Press.

Lakatos, Imre (1970) *Criticism and the Growth of Knowledge*. Eds. Imre Lakatos and Alan Musgrave. Cambridge: Cambridge University Press.

—— (1978) *The Methodology of Scientific Research Programmes: philosophical papers*. Volume I. Eds. John Worrall and Gregory Currie. Cambridge: Cambridge University Press.

Latour, Bruno (1987) *Science in Action*. Cambridge: Cambridge University Press.

Laudan, Larry (1981) 'A Confutation of Convergent Realism' *Philosophy of Science* 48: 19–49.

—— (1990) 'Demystifying Underdetermination' *Minnesota Studies in the Philosophy of Science* 14: 267–97.

Layne, Christopher (1994) 'Kant or Cant? The Myth of Democratic Peace' *International Security* 19: 19–49.

Lebow, Richard Ned (1981) *Between Peace and War: the nature of international crisis*. Baltimore, Md.: Johns Hopkins University Press.

—— (2003) *The Tragic Vision of Politics: ethics, interests and orders*. Cambridge: Cambridge University Press. .

Lehrer, Keith (1986) 'Chisholm on Certainty'. In *Roderick Chisholm*. Ed. Radu Bogdan. Dordrecht: Reidel, pp. 157–67.

Leibniz, G.W.F. (1956) 'Letters to Clarke'. In *The Leibniz–Clarke Correspondence*. Ed. H.G. Alexander. Manchester: Manchester University Press, pp. 5–126

Leplin, Jerrett (1997) *A Novel Defense of Scientific Realism*. New York: Oxford University Press.

Le Roy, Edouard (1901) 'Un positivisme nouveau' *Revue de Métaphysique et de Morale* 9: 138–53.

Levy, Jack (1989) 'The Causes of War: a review of theories and evidence'. In *Behavior, Society, and Nuclear War*. Eds. P.E. Tetlock et al. New York: Oxford University Press.

Lewis, Bernard (1990) *Race and Slavery in the Middle East: an historical enquiry*. New York: Oxford University Press.

—— (2002) *What Went Wrong? The Western Impact and Middle Eastern Response*. Oxford: Oxford University Press.

Linklater, Andrew (1990) *Beyond Realism and Marxism*. London: Macmillan.

Little, Daniel (1991) *Varieties of Social Explanation*. Boulder, Col.: Westview Press.

—— (1993a) 'On the Scope and Limits of Generalizations in the Social Sciences' *Synthèse* 97: 183–207.

—— (1993b) 'Evidence and Objectivity in the Social Sciences' *Social Research* 60: 363–96.

—— (1994) 'Microfoundations of Marxism'. In *Readings in the Philosophy of the Social Sciences*. Ed. M. Martin and L. Macintyre. Cambridge, Mass.: MIT Press.

—— (1998) *Microfoundations, Method, and Causation*. New Brunswick, NJ: Transaction Publishers.

Little, Richard (2000) 'The English School's Contribution to the Study of International Relations' *European Journal of International Relations* 6: 395–422.

Lyotard, François (1993) *Political Writings*. Trans. Bill Reading with Kevin Paul Geiman. Minneapolis, Minn.: University of Minnesota Press.

MacIntyre, Alastair (1973) 'The Essential Contestability of Some Social Concepts' *Ethics* 84: 1–9.

Majer, Ulrich (1995) 'Geometry, Intuition and Experience from Kant to Husserl' *Erkenntnis* 42: 261–85.

Mann, Michael (2001) 'Military Professionalism and the Democratic Peace: how German is it?' In *Democracy, Liberalism, and War: rethinking the democratic peace debate*. Eds. T. Barkawi and M. Laffey. Boulder, Col.: Lynne Rienner, pp. 67–86.

Maoz, Zeev and Nazrin Abdolali (1989) 'Regime Types and International Conflict' *Journal of Conflict Resolution* 33: 3–35.

Maoz, Zeev and Bruce Russett (1993) 'Normative and Structural Causes of Democratic Peace, 1946–1986' *American Political Science Review* 87: 624–38.

Masterman, Margaret (1974) 'The Nature of a Paradigm'. In *Criticism and the Growth of Knowledge*. Eds. Imre Lakatos and Alan Musgrave. Cambridge: Cambridge University Press, pp. 59–89.

McAllister, James W. (1993) 'Scientific Realism and the Criteria of Theory-Choice' *Erkenntnis* 38: 203–22.

Mearsheimer, John (1994–5) 'The False Promise of International Institutions' *International Security* 19: 5–49.

Mearsheimer, John and Steven M. Walt (2003) 'An Unnecessary War' *Foreign Policy* 134: 51–9.

Metz, Steven (2003–4) 'Insurgency and Counterinsurgency in Iraq' *The Washington Quarterly* 27: 25–37.

Midlarsky, Manus (1992) 'The Origins of Democracy in Agrarian Society' *Journal of Conflict Resolution* 36: 454–77.

—— (1995) 'Environmental Influences on Democracy: aridity, warfare, and a reversal of the causal arrow' *Journal of Conflict Resolution* 39: 224–62.

Mill, John Stuart (1974) *A System of Logic: ratiocinative and inductive*. Toronto: University of Toronto Press.

Modelski, George (1987) *Long Cycles in World Politics*. Seattle, Wash.: University of Washington Press.

Modelski, George and Kazimierz Poznanski (1996) *International Studies Quarterly – Special Issue: evolutionary paradigms in the social sciences* 40 (September).

Modelski, George and William R. Thompson (1989) 'Long Cycles and Global War'. In *Handbook of War Studies*. Ed. Manus I. Midlarsky. Boston, Mass.: Unwin Hyman, pp. 1–23.

Morgenthau, Hans J. (1949) *Politics Among Nations: the struggle for power and peace*. New York: Alfred A. Knopf.

Mufson, Stephen and Thomas E. Ricks (2001) 'Debate over Targets Highlights Difficulty of War on Terrorism' *New York Times*, 21 September, A25.

Neurath, Otto (1939) *Modern Man in the Making*. New York and London: Alfred A. Knopf.

Niebuhr, Reinhold (1940) *Christianity and Power Politics*. New York: Charles Scribners' Sons.

North, Douglass C. (1987) 'Institutions and Economic Growth: an historical introduction'. Paper prepared for the Conference on Knowledge and Institutional Change. University of Minnesota. (November).

Olson, Mancur (1965) *The Logic of Collective Action: public goods and the theory of groups*. Cambridge, Mass.: Harvard University Press.

Oneal, John R. and Bruce Russett (1997) 'The Classical Liberals Were Right: democracy, interdependence, and conflict, 1950–1985' *International Studies Quarterly* 41: 267–94.

—— (1999a) 'Is the Liberal Peace Just an Artifact of Cold War Interests? Assessing recent critiques' *International Interactions* 25: 1–29.

—— (1999b) 'The Kantian Peace: the pacific benefits of democracy, interdependence, and international organizations, 1885–1992' *World Politics* 52: 1–37.

—— (2001) 'Clear and Clean: the fixed effects of democracy and economic interdependence' *International Organization* 55: 469–85.

Onuf, Nicholas Greenwood (1989) *World of Our Making: rules and rule in social theory in international relations*. Columbia, SC: University of South Carolina Press.

Organski, A.F.K. (1958) *World Politics*. New York: Knopf.

Organski, A.F.K. and Jacek Kugler (1980) *The War Ledger*. Chicago, Ill.: University of Chicago Press.

Owen, John (1994) 'How Liberalism Produces Democratic Peace' *International Security* 9: 87–125.

—— (1997) *Liberal Peace, Liberal War, American Politics and International Security*. Ithaca, NY: Cornell University Press.

Parsons, Talcott (1937) *The Structure of Social Action: a study in social theory with special reference to a group of recent European writers*. New York: McGraw-Hill.

Patomäki, Heikki (2002) *After International Relations: critical realism and the (re)construction of world politics*. London: Routledge.

Patomäki, Heikki and Colin Wight (2000) 'After Postpositivism: the promises of critical realism' *International Studies Quarterly* 44: 213–37.

Paul, T.V. (1994) *Assymetric Conflicts*. Cambridge: Cambridge University Press.

Peirce, Charles Sanders (1932) *The Collected Papers of Charles Sanders Peirce*. Eds. Charles Hartshorne and Paul Weiss, vol. II, *Elements of Logic*. Cambridge, Mass.: Harvard University Press.

Poincaré, Henri (1905) *Science and Hypothesis*. Trans. George Bruce Halstead. Intro. Josiah Royce. New York: Science Press.

—— (1914) *Science and Method*. Trans. Francis Maitland. London: T. Nelson & Sons.

—— (1952) *Science and Hypothesis*. Trans. W.J. Greenstreet. New York: Dover.

Pollack, Kenneth M. (2002) 'Next Stop Baghdad?' *Foreign Affairs* (81) 2: 32–47.

—— (2003) 'Securing the Gulf' *Foreign Affairs*. (82) 4: 2–16.

Pollins, Brian M. (forthcoming) 'Beyond Logical Positivism: reframing King, Keohane and Verba's designing social enquiry'. In *Theory and Evidence*. Eds. R. Ned Lebow and Friedrich Kratochwil.

Popper, Karl R. (1945) *The Poverty of Historicism*. London: G. Routledge & Sons.

—— (1948) 'Prediction and Prophecy, and Their Significance for Social Theory'. *Proceedings of the Xth International Congress of Philosophy*. Amsterdam: North-Holland.

—— (1957) 'The Propensity Interpretation of the Calculus of Probability and the Quantum Theory'. In *The Colston Papers*. Ed. Stephan Körner. London: Butterworth Scientific Publications, pp. 57–60.

—— (1968) *The Logic of Scientific Discovery*, 2nd edn. New York: Harper Torchbooks.

—— (1971a) *The Open Society and Its Enemies. 1 The spell of Plato*. Princeton, NJ: Princeton University Press.

—— (1971b) *The Open Society and Its Enemies. 2 Hegel and Marx*. Princeton, NJ: Princeton University Press.

Prior, Arthur N. (1962) *Formal Logic*. Oxford: Clarendon Press.

Puchala, Donald (1991) 'Woe to the Orphans of the Scientific Revolution'. In *The Evolution of Theory in International Relations*. Ed. Robert L. Rothstein. Columbia, SC: University of South Carolina Press, pp. 39–60.

Putnam, Hilary (1975) *Philosophical Papers*. New York: Cambridge University Press.

Quester, George H. (2002) *Before and After the Cold War: using past forecasts to predict the future*. London: Frank Cass Publishers.

Quine, Willard van Orman (1953) 'Two Dogmas of Empiricism'. In *From a Logical Point of View*. Cambridge, Mass.: Harvard University Press, pp. 20–46.

—— (1990) *Pursuit of Truth*. Cambridge, Mass.: Harvard University Press.

Quine, Willard van Orman and J.S. Ullian (1978) *The Web of Belief*. New York: Random House.

Ragin, Charles C. (1988) Comparative Method: moving beyond qualitative and quantitative strategies. Berkeley, Calif.: University of California Press.

Ray, James Lee (1995) *Democracy and International Politics: an evaluation of the democratic peace proposition*. Columbia, SC: University of South Carolina Press.

Ray, James Lee and Bruce Russett (1996) 'The Future as Arbiter of Theoretical Controversies: predictions, explanations and the end of the Cold War' *British Journal of Political Science* 25: 441–70.

Raymond, Gregory A. (1994) 'Democracy, Disputes, and Third Party Intermediaries' *Journal of Conflict Resolution* 38: 24–42.

—— (1996) 'Demosthenes and Democracies: regime-types and arbitration outcomes' *International Interactions* 22: 1–20.

Reichenbach, Hans J. (1936) 'Logistic Empiricism in Germany and the Present State of its Problems' *Journal of Philosophy* 33: 141–60.

—— (1938) *Experience and Prediction*. Chicago, Ill.: University of Chicago.

—— (1958) *The Philosophy of Space-Time*. Trans. Maria Reichenbach and John Freund. New York: Dover.

Richardson, Lewis F. (1960) *The Statistics of Deadly Quarrels*. Pittsburgh, Pa.: Boxwood.

Rosenau, James (1980) *The Scientific Study of International Relations*, rev. edn. London: Pinter.

Rosenkranz, Gary (1981) 'The Nature of Geometry' *American Philosophical Quarterly* (April): 101–10.

Rousseau, David , Christopher Gelpi, Dan Reiter and Paul Huth (1996) 'Assessing the Dyadic Nature of the Democratic Peace, 1918–1988' *American Political Science Review* 90: 512–33.

Ruggie, John Gerard (1986) 'Continuity and Transformation in the World Polity: towards a neorealist synthesis'. In *Neorealism and Its Critics*. Ed. R. Keohane. New York: Columbia University Press, pp. 131–57.

——(1996) *Winning the Peace: America and the world order in the new era*. New York: Columbia University Press.

Rummel, Rudolf J. (1968) 'Domestic Attributes and Foreign Conflict'. In *Quantitative International Politics: insights and evidence*. Ed. J. David Singer. New York: Free Press, pp. 187–214.

—— (1979) *War, Power, Peace*. Vol. IV of *Understanding Conflict and War*. Beverly Hills, Calif.: Sage.

—— (1981) *The Just Peace*. Vol. V of *Understanding Conflict and War*. Beverly Hills, Calif.: Sage.

—— (1983) 'Libertarianism and International Violence' *Journal of Conflict Resolution* 27: 27–71.

—— (1992) *Democide*: *Nazi genocide and mass murder*. New Brunswick, NJ: Transaction Publishers.

—— (1997) *Power Kills: democracy as a method of nonviolence*. New Brunswick, NJ: Transaction Publishers.

Rupert, Mark (2001) 'Democracy, Peace: What's not to love?' In *Democracy, Liberalism and War: rethinking the democratic peace debate*. Eds. T. Barkawi and M. Laffey. Boulder, Col.: Lynne Rienner, pp. 153-72.

Russell, Bertrand (1897) *Essay on the Foundations of Geometry*. Cambridge: Cambridge University Press.

—— (1918) 'On the Notion of Cause'. In *Mysticism and Logic and Other Essays*. New York: Longmans, pp. 132–51.

Russett, Bruce M. (1963) 'The Calculus of Deterrence' *Journal of Conflict Resolution* 7: 97–109.

—— (1985) 'The Mysterious Case of Vanishing Hegemony; or, is Mark Twain really dead?' *International Organization* 39: 207–31.

—— (1990) *Controlling the Sound: the democratic governance of national security*. Cambridge, Mass.: Harvard University Press.

—— (1993) *Grasping the Democratic Peace: principles for a post-cold war world*. Princeton, NJ: Princeton University Press.

—— (1995) 'The Democratic Peace: and yet it moves' *International Security* 19: 164–75.

—— (1996) 'Counterfactuals about War and Its Absence.' In *Counterfactual Thought Experiments in World Politics: logical, methodological, and psychological perspectives*. Eds. Philip E. Tetlock and Aaron Belkin. Princeton, NJ: Princeton University Press, pp. 171–186.

Russett, Bruce M. and John R. Oneal (2001) *Triangulating Peace: democracy, trade and international organization*. New York: W. W. Norton & Co. Inc.

Ryle, Gilbert (1949) *The Concept of Mind*. London: Hutchinson's University Library.

Salmon, Wesley C. (1979) *Hans Reichenbach: logical empiricist*. Dordrecht: Reidel.

—— (1980) 'Probabilistic Causality' *Pacific Philosophical Quarterly* 61: 51–74.

—— (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

—— (1989) 'Four Decades of Scientific Explanation'. In *Minnesota Studies in the Philosophy of Science*, Vol. XIII *Scientific Exploration*. Eds. P. Kitcher and W.C. Salmon. Minneapolis, Minn.: University of Minnesota Press, pp. 3–219.

Sankey, Howard (1994) *The Incommensurability Thesis*. Avebury: Brookfield.

Schlick, Moritz (1985) *General Theory of Knowledge*. Trans. Albert E. Blumberg. La Salle, Ill.: Open Court.

Schrödinger, Erwin (1935) *Science and the Human Temperament*. Trans. J. Murphy and W.H. Johnston. New York: W.W. Norton & Co.

Schutz, Albert (1967) *The Problem of Social Reality*. Edited and introduction by M. Natanson. The Hague: M. Nijhoff.

Schweller, Randall L. (1997) 'New Realist Research on Alliances: refining, not refuting, Waltz's balancing proposition' *American Political Science Review* 91: 927–30.

Scriven, Michael (1959) 'The Logic of Criteria' *Journal of Philosophy* 56: 857–67.

—— (1962) 'Explanations, Predictions and Laws'. In *Minnesota Studies in the Philosophy of Science*, Vol. III *Scientific Explanation, Space, and Time*. Eds. Herbert Feigl and Grover Maxwell. Minneapolis, Minn.: University of Minnesota Press, pp. 170–230.

—— (1975) 'Causation as Explanation' *Noûs* 9: 3–16.

Sellars, Wilfred S. (1956) 'The Myth of the Given: Three Lectures on Empiricism and the Philosophy of Mind', Lectures delivered 1, 8, 15 March 1956 at University of London. Published as, 'Empiricism and the Philosophy of Mind'. In Minnesota Studies in the Philosophy of Science Vol. I. The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Edited by Herbert Feigl and Michael Scriven. (Minneapolis: University of Minnesota Press): 253-329.

—— (1963) 'Empiricism and the Philosophy of Mind'. In *Science, Perception and Reality*. London: Routledge & Kegan Paul, pp. 127–96.

—— (1965) 'Scientific Realism or Irenic Instrumentalism'. In *Boston Studies in the Philosophy of Science*, Vol. II. New York: Humanities Press, pp. 171–204.

Shapiro, Ian and Alexander Wendt (1992) 'The Difference that Scientific Realism Makes: social science' *Politics and Society* 20: 197–223.

Sidelle, Alan (1989) *Necessity, Essence and Individuation: a defense of conventionalism*. Ithaca, NY: Cornell University Press.

Singer, J. David (1961) 'The Levels of Analysis Problem in International Relations'. In *The International System*. Eds. Klaus Knorr and Sidney Verba. Princeton, NJ: Princeton University Press, pp. 77–92.

Singer, J. David and Melvin Small (1972) 'The Correlates of War Project: an interim report and rationale' *World Politics* 24: 243–70.

Skocpol, Theda (1992) *Protecting Soldiers and Mothers: the political origins of social policy in the United States*. Cambridge, Mass.: Harvard University Press.

Small, Melvin and J. David Singer (1976) 'The War-Proneness of Democratic Regimes' *Jerusalem Journal of International Relations* 1: 50–69.

—— (1982) *Resort to Arms: international and civil wars 1816–1929*. Los Angeles, Calif.: Sage.

Smith, Steve (1996) 'Positivism and Beyond'. In *International Theory: positivism and beyond*. Eds. Steve Smith, Ken Booth and Maria Zalewski. Cambridge: Cambridge University Press, pp. 11–44.

Snyder, Jack and Thomas Christensen (1997) 'Progressive Research on Degenerative Alliances' *American Political Science Review* 91: 919–22.

Sober, Elliot (1988) *Reconstructing the Past: parsimony, evolution and inference*. Cambridge, Mass.: MIT Press.

Sorenson, Theodore (1965) *Kennedy*. New York: Harper & Row.

Spiro, David E. (1994) 'The Insignificance of the Liberal Peace' *International Security* 19: 50–86.

Starr, Harvey (1997) 'Democracy and Integration: why democracies don't fight each other' *Journal of Peace Research* 34: 153–62.

Steinbruner, John (1974) *Cybernetic Theory of Decision*. Princeton, NJ: Princeton University Press.

Stephenson, June (1998) *Poisonous Power: childhood roots of tyranny*. Palm Desert, CA: Diemer, Smith Publishing.

Suganami, Hidemi (1996) *On the Causes of War*. Oxford: Oxford University Press.

—— (2003) 'Beyond the English School'. Paper presented ISA Portland March.

Sun Tzu (1971) *The Art of War*. Trans. Samuel B. Griffith. Oxford: Oxford University Press.

Taylor, Charles (1985) 'Interpretation and the Sciences of Man'. In *Philosophy and the Human Sciences: philosophical papers*, Vol. II. Cambridge: Cambridge University Press, pp. 15–57.

Thayer, Bradley A. (2001) 'Correspondence: starting the evolution without us' *International Security* 26: 194–8.

Thompson, William R. (1996) 'Democracy and Peace: putting the cart before the horse?' *International Organization* 50: 141–74.

Thompson, William R. and Richard Tucker (1997) 'A Tale of Two Democratic Peace Critiques' *Journal of Conflict Resolutions* 41: 428–54.

Tickner, J. Ann (2001) *Gendering World Politics: issues and approaches in in the post-Cold War era*. New York: Columbia University Press.

Tilly, Charles (2003) *Politics of Coercive Violence*. Cambridge: Cambridge University Press.

Tures, John A. (2001) 'Democracies as Intervening States: a critique of Kegley and Hermann' *Journal of Peace Research* 38: 227–35.

Vaitkus, Steven (1994) 'The Realist Image in Social Science, or a Realist's Categorization of Social Thinking' *Philosophy of the Social Sciences* 24: 76–83.

van Fraassen, Bas C. (1980) *The Scientific Image*. New York: Oxford University Press.

—— (1989) *Laws and Symmetry*. Oxford: Clarendon.

Van Heijenoort, Jean (1966) *Frege to Gödel: sourcebook on mathematical logic*. Cambridge, Mass.: Harvard University Press.

Vasquez, John A. (1997) 'The Realist Paradigm and Degenerative Versus Progressive Research Programs: an appraisal of neotraditional research on Waltz's balancing proposition' *American Political Science Review* 91: 899–912.

—— (1998) *The Power of Power Politics: from classical realism to neotraditionalism*. Cambridge and New York: Cambridge University Press.

—— (2002) 'The New Debate on Balancing: A Reply to my critics'. In *Realism and the Balancing of Power: A new debate*. Eds. J. Vasquez and C. Elman. Upper Saddle River, NJ: Prentice Hall.

Von Mises, Richard (1939) *Probability, Statistics and Truth*. Trans. J. Neyman, D. Sholl and E. Rabinowitsch. New York: Macmillan.

Von Wright, G.H. (1971) *Exploration and Understanding*. Ithaca, NY: Cornell University Press.

Waldman, Peter (2004) 'A Historian's Take on Islam Steers U.S. in Terrorism Fight: Bernard Lewis's blueprint—sowing Arab democracy—is facing a test in Iraq' *Wall Street Journal* 3 February 2004: 1.

Walker, R.J.B. (1993) *Inside/Outside*. Cambridge: Cambridge University Press.

Wallace, Michael D. (1982) 'Arms and Escalation: two competing hypotheses' *International Studies Quarterly* 26: 37–56.

Walt, Steven M. (1997) 'The Progressive Power of Realism' *American Political Science Review* 91: 931–5.

Waltz, Kenneth N. (1959) *Man, the State and War*. New York: Columbia University Press.

—— (1979) *Theory of International Politics*. Reading, Mass.: Addison-Wesley.

—— (1986) 'Reflections on *Theory of International Politics*: a response to my critics'. In *Neorealism and Its Critics*. Ed. Robert O. Keohane. New York: Columbia University Press, pp. 322–5.

—— (1997) 'Evaluating Theories' *American Political Science Review* 91: 913–17.

Watson, Adam (1992) *The Evolution of International Society: a comparative historical analysis*. London: Routledge.

Wayman, Frank W. and Paul F. Diehl (1994) 'Realism Reconsidered'. In *Reconstructing Realpolitik*. Eds. F. Wayman and P. Diehl. Ann Arbor, Mich.: University of Michigan Press, pp. 3–26.

Weart, Spencer R. (1994) 'Peace Among Democratic and Oligarchic Republics' *Journal of Peace Research* 31: 299–316.

Weber, Katja (2000) *Hierarchy Amidst Anarchy: transaction costs and institutional choice*. Albany, NY: SUNY Press.

Weber, Max (1949) *The Methodology of the Social Sciences*. New York: Free Press.

—— (1974) 'Subjectivism and Determinism: a translation of Weber, Roescher and Knies und das irrationalitatproblem.' In *Positivism and Sociology*. Ed. A. Giddens. London: Heinemann, pp. 23–31.

Weede, Erich (1984) 'Democracy and War Involvement' *Journal of Conflict Resolution* 28: 649–64.

Wendt, Alexander (1987) 'The Agent–Structure Problem in International Relations Theory' *International Organization* 41: 335–70.

—— (1992) 'Anarchy is What States Make of It: the social construction of power politics' *International Organization* 46: 391–425.

—— (1995) 'Constructing International Politics' *International Security* 20: 71–81.

—— (1999) *Social Theory of International Politics*. Cambridge: Cambridge University Press.

Wiener, Norbert (1949) *Cybernetics*. New York: John Wiley & Sons.

Wight, Colin (1996) 'Incommensurability and Cross-Paradigm Communications in International Relations Theory: what's the frequency Kenneth?' *Millennium: Journal of International Studies* 25: 291–319.

Wight, Martin (1991) *International Theory: three traditions*. Eds. G. Wight and B. Porter. Leicester: University of Leicester Press.

Williams, Michael C. and Keith Krause (1997) 'Introduction'. In *Critical Security: concepts and cases*. Eds. Michael C. Williams and Keith Krause. Minneapolis, Minn.: University of Minnesota Press.

Winch, Peter (1958) *The Idea of a Social Science*. London: Routledge & Kegan Paul.

Wittgenstein, Ludwig (1953) *Philosophical Investigations*. Tr. G.E.M. Anscombe. Oxford: Basil Blackwell.

Woodward, Bob (2002) *Bush at War*. New York: Simon & Schuster.

Wyn-Jones, Richard (1999) *Security, Strategy, and Critical Theory*. Boulder, Col.: Lynne Rienner.

Young, Oran (1982) 'Regimes Dynamics: the rise and fall of international regimes' *International Organization* 36: 277–97.

Zehfuss, Maja (2002) *Constructivism in International Relations: the politics of reality*. Cambridge: Cambridge University Press.

Index