
The Future's Back
Nuclear Rivalry, Deterrence Theory, and Crisis Stability
after the Cold War

Given the likely proliferation of weapons of mass destruction in the future, and the failure of “non-proliferation” regimes to provide effective barriers to the spread of nuclear technology in the past, efforts towards greater understanding of nuclear rivalry are crucial. *The Future's Back* provides a valuable framework for organizing and evaluating research on superpower rivalry and nuclear deterrence.

Arguing that previous critiques of rational choice and deterrence theory are unconvincing, Frank Harvey constructs a new set of empirical tests of rational deterrence theory to illuminate patterns of interaction between rival nuclear powers. He analyses the crisis management techniques used by the United States and the Soviet Union in twenty-eight post-war crises and isolates factors that promote or inhibit escalation of these crises. This “crises”-based data set serves as a basis for identifying patterns of response when one nuclear state is threatened by another.

The Future's Back offers new directions for testing that emphasize a more unified approach to theory building and assesses the feasibility of alternative courses of action to prevent escalation of future disputes characterized by nuclear rivalry.

FRANK P. HARVEY is associate professor of political science, Dalhousie University.

This page intentionally left blank

The Future's Back

*Nuclear Rivalry, Deterrence Theory,
and Crisis Stability after
the Cold War*

FRANK P. HARVEY

McGill-Queen's University Press
Montreal & Kingston • London • Buffalo

© McGill-Queen's University Press 1997
ISBN 0-7735-1605-0 (cloth)
ISBN 0-7735-1606-9 (paper)

Legal deposit third quarter 1997
Bibliothèque nationale du Québec

Printed in Canada on acid-free paper

This book has been published with the help of a grant from the Humanities and Social Sciences Federation of Canada, using funds provided by the Social Sciences and Humanities Research Council of Canada.

McGill-Queen's University Press acknowledges the support received for its publishing program from the Canada Council's Block Grants program.

Canadian Cataloguing in Publication Data

Harvey, Frank P. (Frank Paul), 1962-

The future's back : nuclear rivalry, deterrence theory, and crisis stability after the Cold War

Includes bibliographical references and index.

ISBN 0-7735-1605-0 (bound) -

ISBN 0-7735-1606-9 (pbk.)

1. Nuclear crisis stability. 2. Nuclear crisis control.
3. Deterrence (Strategy). 4. Security, International.
- I. Title.

U264.H368 1997

355.02'17

C97-900545-0

Typeset in New Baskerville 10/12
by Caractéra inc., Québec City

To Kalli G

This page intentionally left blank

Contents

Tables and Figures	ix
Acknowledgments	xi
Introduction	xiii
1 Cumulation in International Relations: Methodological Constraints and Solutions	3
2 Nuclear Deterrence: The Record of Aggregate Testing	19
3 New Directions for Aggregate Testing	47
4 Game Theory and Superpower Rivalry, 1948–88	56
5 U.S.-Soviet Rivalry and Nuclear Deterrence	94
6 Nuclear Stability after the Cold War	112
7 Conventional Deterrence and Compellence Theory: Perspectives on Testing after the Cold War	121
Notes	145
Bibliography	171
Index	191

This page intentionally left blank

Tables and Figures

T A B L E S

- 2.1 Classification of Deterrence Literature 21
- 2.2 Nuclear Constraints and Conflict Type 32
- 2.3 Classes of Deterrence/Compellence Encounters 34
- 2.4 Coefficients of Correspondence 37
- 3.1 Superpower Rivalry in Foreign Policy Crises,
1948-88 49
- 4.1 Intermediate Payoff Combinations and Resulting Error
Levels 70
- 4.2 Superpower Bargaining Techniques: Provocation and
Retaliation 73
- 4.3 Trigger and Response in Foreign Policy Crises,
1948-88 74
- 5.1 Degrees of Geostrategic Importance 98
- 5.2 Degrees of Heterogeneity 98
- 5.3 Issues 99

x Tables and Figures

5.4	Gravity of Threats to Superpower Values	100
5.5	Perceived Threats to Superpower Influence	101
5.6	Crisis Triggers	101
5.7	Major Responses	103
5.8	Crisis-Management Techniques	103
5.9	Intensity of Violence	104
5.10	Centrality of Violence	104
5.11	Provocation and Retaliation in Foreign Policy Crises, 1948–88	105
7.1	Empirical Domain of Immediate Deterrence/ Compellence Tests	124
7.2	Summary of Exchanges with Bosnian Serbs	127
7.3	Consistent Response Sets	129
7.4	Inconsistent Response Sets	130
7.5	Consistent Response Sets: Bosnia, 1993–95	132

FIGURES

1.1	Hypothetical profiles of foreign policy elites	11
1.2	Rationality and stress in foreign policy crises	15
4.1	The threat game	60
4.2	Retaliation functions for TFT, OT, GRIT, and GRIM	65
4.3	Provocation and retaliation in superpower crises	75
4.4	Margin of predictive error, OT vs TFT	78
4.5	Margin of predictive error, OT vs TFT	79
4.6	Optimal threats and satisfaction	83
4.7	Satisfaction and goal achievement, GRIM vs GRIT	84

Acknowledgments

I would like to thank the Canadian Institute for International Peace and Security, the Cooperative Security Competitions Program (Department of External Affairs and Trade, grant no. CS008-2), and the Social Science and Humanities Research Council of Canada (grant no. 410-93-0574) for the generous financial assistance I received to complete this project.

I have benefited enormously from the help and support of Patrick James. Every significant academic achievement in my life, including this book, would have been impossible without his wisdom and guidance. To him I owe my career. I am indebted to Michael Brecher for his help and advice over the years and for comments and suggestions made on earlier drafts of this manuscript.

I would like to thank Anupa, the most important person in my life, for her infinite patience. To her I owe my happiness. The greatest debt of all is owed to my family for their constant encouragement.

This page intentionally left blank

Introduction

The disintegration of the Soviet Union heralded the most significant transformation in global politics since World War II, a turning point that was destined to have a profound impact on a system which, until recently, was considered by many to be relatively stable. Forces that guided relations between East and West dissolved as the collapse of a superpower produced a new environment where opportunities for conflict and cooperation became less clearly defined.

Although the significance of this transformation is widely recognized, consensus breaks down on the precise implications for nuclear proliferation, nuclear deterrence, and superpower crisis management. Should the decline and fall of the Soviet Union provoke optimistic or pessimistic forecasts about global peace and security as we approach the twenty-first century? Will the shift away from "power bipolarity" create a more stable or a more hostile nuclear environment? Are these changes likely to intensify or diminish the horizontal proliferation of nuclear weapons in the Middle East and other parts of the developing world? If proliferation intensifies, will changes in the Cold War balance affect escalation and management of military-security crises between nuclear states in the next decade? In sum, will we come to miss the relative stability of the "long postwar peace" (Gaddis 1986, 1987, 1991; Mearsheimer 1990a,b,c; Kegley 1991) or welcome the relative security of the "New World Order," with the United States occupying the sole superpower role? Providing answers to these questions is an important challenge to international relations (IR) scholarship today.

These questions are the focus of this inquiry. The title, *The Future's Back*, refers to the importance of speculating about the impact of change, based on the knowledge assembled to explain interstate relations in the past.¹ Central to any analysis of change is the idea of comparison; something changes in comparison to a previous state (Dunn 1981). Consequently, the objective of this investigation is threefold: (1) to examine the crisis management techniques of the United States and the Soviet Union during intense crises between 1948 and 1988, (2) to isolate factors that promoted or inhibited escalation of these crises, and (3) to discover whether similar forces present comparable opportunities and constraints for potential nuclear rivals in the future.

Those who probe these questions continue to cultivate their positions with reference, implicitly or explicitly, to preferred theories about interstate relations, international conflict, and crisis management. Generated from decades of conflict research, these theories, models, and frameworks serve as signposts towards one or another solution to the puzzle about future nuclear stability. It would appear, then, that a comprehensive evaluation of theoretical accomplishments is the best approach to developing a solid and generally accepted assessment of global change. As Mearsheimer (1990a, 9) recommends, we simply "decide which theories best explain the past, and most directly apply to the future, and then employ these theories to explore the consequences of possible scenarios." Accepting this challenge, however, will not be easy, for there is no theoretical consensus in IR – or in conflict theory in particular.

If only the "hard" sciences were cumulative and progressive, then Mearsheimer's challenge would be chimerical – there would be no generally accepted social-scientific theory or body of knowledge upon which to build.² Thus the problem, and the main focus of this project, is this: as a community of scholars, we have yet to develop a method of cumulating our theoretical insights. Notwithstanding the tremendous amount of time and energy spent on developing conflict theory, there remains no clear evidence of cumulative knowledge, a prerequisite for investigating the consequences of change. As a result, we are not well equipped to provide persuasive answers to many of the most pressing questions of the day.

This project attempts to respond to the challenge by developing a more focused approach towards cumulation. This task will be accomplished by evaluating and synthesizing what has been discovered about intense U.S.-Soviet crisis management techniques, which in turn will serve as the basis for predictions about the impact of recent changes on global nuclear stability. Whether these changes have negative or positive implications, it will be argued, depends largely on the factors

that have promoted or inhibited violent superpower conflict since 1948. If the relative stability of Cold War nuclear crises can be traced to the behavioural dynamics embedded within U.S.-Soviet rivalry – an assertion to be developed and tested more fully in subsequent chapters – then the current transformation the system is undergoing may produce a nuclear environment substantially less stable than the system of 1945–88, notwithstanding the absence of U.S.-Soviet tension.

OUTLINE OF THE ARGUMENT

This investigation comprises seven chapters.³ Chapter 1 focuses on the development and cumulation of theory in IR scholarship. The findings suggest that the relative impact of different barriers to cumulation is far from uniform and that methodological impediments arising from the sustained disagreement over the proper nature of inquiry offer the most compelling explanation for the lack of progress. Reflecting these debates, the field has become saturated with efforts to falsify potentially important theories (and paths to knowledge) by emphasizing the limitations inherent in a given approach. In the conviction that these disagreements are counterproductive, chapter 1 goes on to explore the research program in *crisis management* and *deterrence*. The purpose is twofold: (1) to probe for information about superpower rivalry and nuclear deterrence, and (2) to illustrate the counterproductive nature of debates over methods, in which the rational choice and political psychology schools represent the two extremes of crisis management theory. The chapter concludes with a framework for organizing and evaluating research on superpower rivalry and nuclear deterrence that encourages recognition of cumulation and consensus, while offering new avenues for testing (developed more formally in chapters 3, 4, and 5).

Chapter 2 applies the framework to U.S.-Soviet crisis behaviour and nuclear deterrence. Research on the subject is primarily speculative and focuses on the behaviour expected from antagonists who are able to threaten each other with unacceptable damage. By comparison, little attention has been directed towards examining the relationship between expected behaviour and actual results during military-security crises. The most significant problem with traditional analyses is the paucity of aggregate investigation. In particular, it is still unclear whether states act according to the logic derived from standard applications of nuclear deterrence. The implications with respect to designing policy to deal with nuclear proliferation are obvious.

Although the deterrent role played by nuclear weapons cannot be tested directly in the absence of nuclear war, the role of these forces

can be assessed indirectly by focusing on attendant propositions embedded within nuclear deterrence theory. Chapter 3 describes two alternative approaches to aggregate testing that build on previous research and specifies the coding procedures and selection criteria for the set of superpower crises during 1948–88 that serve as the empirical base for the aggregate analysis. Formal presentation of the tests appears in chapters 4 and 5.

Research on the bargaining behaviour of nuclear powers in the past is expected to facilitate greater comprehension of potential future interaction involving states with nuclear weapons. With this in mind, chapter 6 discusses the implications of the findings for post-Cold War nuclear stability and assesses the likely impact of nuclear proliferation on global peace and security, offering policy recommendations along the way. Chapter 7 shifts focus to *conventional* deterrence, offering a few thoughts on new directions for testing in this realm and concluding with observations about the enduring relevance of realist theories like deterrence theory.

The Future's Back

This page intentionally left blank

1 Cumulation in International Relations: Methodological Constraints and Solutions

The sustained disagreement over the nature of inquiry and related debates about appropriate criteria for judging progress are significant impediments to cumulation – that is, to the development of a generally accepted body of knowledge – in the subfield of international relations.¹ While some commentators believe that knowledge is advanced by in-depth study of a few crucial cases, others are convinced that aggregate analysis of many cases is required for valid generalizations (Bueno de Mesquita 1985). A second disagreement continues between adherents of deductive and inductive approaches to explanation; a third concerns quantitative analysis as opposed to more traditional qualitative methods; and a final division exists among those who subscribe to system-level (macro) as opposed to actor-level (micro) analyses.² Although all four are important, only the inductive/deductive debate will be described here, in order to reveal the effect these kinds of disputes have on cumulative knowledge about U.S.-Soviet rivalry, crisis management, and deterrence theory.

Deduction involves reasoning from general laws, that is axioms, to specific instances. Hypotheses are subsequently derived and tested on the basis of empirical evidence. The axioms, which combine to create models or theories, are products of intuition, common sense, simple observations, or a credible interpretation of history. The inductive approach, on the other hand, involves a different route to knowledge: generalizing from experience and observed patterns. Propositions that purport to explain those patterns are generated and empirically tested, and on the basis of these tests, which usually entail correlations

between independent and dependent variables (or more sophisticated statistical procedures), attempts are made to generalize by developing a model or theory (that is, the inductive approach is a "bottom up" as opposed to a "top down" approach).

Which approach, deduction or induction, is more conducive to the process of theory building? For proponents of deduction, the most elemental requirement for all research is that hypotheses (or propositions) subjected to verification be logically consistent. In the absence of specification of the logical linkages between the independent and dependent variables of the proposition, the causal (and, by implication, theoretical) linkage between those elements will remain incomplete, regardless of any statistical correlation between variables. A recurrent problem with inductive research is the tendency to assume causation on the basis of correlation and to neglect rigorous hypothesis testing that renders the facts and empirical evidence meaningful.

Two examples from international relations research illustrate the problems associated with testing propositions that are not logically derived from theory: the application of *status inconsistency* to international conflict, and the debate over *system structure* and *stability*.³ Status-inconsistency theory stipulates that leaders of states with greater power, or capabilities, than ascribed status (that is, international recognition or respect) will attempt to address this imbalance through conflict and war, rather than some form of diplomacy. But there is nothing in the logical structure of the arguments to explain why leaders are more prone to vent their frustration over the imbalance through the use of military force or why they would confront the country's low ascribed status by jeopardizing its position further through acts of aggression.

Similar problems apply to much of the work on system structure and stability (Waltz 1964; Deutsch and Singer 1964; Wayman 1984, 1985; Levy 1985; Sabrosky 1985; Brecher and James 1988). Instead of emphasizing correlations between different measures of polarity and stability across time (the traditional testing strategy), researchers should focus on the logical relationship between polarity, capability, and uncertainty (Bueno de Mesquita 1985). The most relevant question is this: Why do various distributions of power create more or less uncertainty, and what are the implications with respect to state behaviour? The classic expositions on polarity and stability are unclear on these points.

Far from criticising the arguments regarding the merits of deduction, others generally accept the importance of rigour in theorizing. The slowly emerging consensus is that if knowledge is the objective, deduction and induction are not mutually exclusive. For one thing, it would be impossible to use an entirely inductive or deductive method

in isolation. Morgenthau's belief (1960) that all international political activity and behaviour is based on the search for, acquisition, and maintenance of power is, in most formulations, a deductive theory of state behaviour. But the axioms and assumptions of the theory are not derived from intuition or common sense alone; they are based on Morgenthau's observations of interactions among states. Similarly, explanatory variables chosen to correlate with one or more dependent variables often serve a specific purpose and are usually based on a tentative theoretical linkage. One can arrive at laws and empirical generalizations through inductive methods and statistical correlations, but one cannot explain those laws: explanation as Waltz (1979) correctly points out, is the function of theory.

Notwithstanding what appears to be the foundation for an emerging consensus, debates over methods persist. Scholars approaching the subject from one perspective frequently disregard the findings from another, simply by virtue of their general approach. This problem is clearly illustrated by a relatively recent symposium in *World Politics* on the issue of deterrence, one closely related to the subject of the present project, strategic response in crises. Because of its almost exclusive focus on methodological issues, the debate as a whole provides an excellent example of a combined effort to brand a theory as invalid by pointing out the limitations inherent in the general approach.

The question addressed was whether or not states behave in accordance with the logic stipulated by rational deterrence theory. Approaching the subject from an inductive perspective, George and Smoke (1989), Jervis (1989), Lebow and Stein (1989a), and other critics cited evidence from comparative case studies that made rational deterrence appear to be an unreliable guide to reality. They argued further that decision makers generally are incapable of functioning as prescribed by rational choice theory, and that assumptions about the rational, unitary decision maker, perfect information, and a politically neutral environment were unrealistic and even harmful: "Rational deterrence theories [pertain to] nonexistent decision makers operating in nonexistent environments" (Lebow and Stein 1989a, 224). Obviously, those who are unconvinced by rational choice models would expect this critique to be especially relevant to cases of presumed nuclear deterrence.

In contrast, Achen and Snidal, relying almost entirely on the axiomatic (that is, deductive) method borrowed from microeconomics and game theory, asserted that proponents of case studies have misinterpreted rational deterrence as a "theory of how decision makers think" (1989, 164). Consequently, comparative case studies provide

weak tests of the validity of rational deterrence theory. Furthermore, the selection of cases by critics is often biased in favour of examining crises in which deterrence has already failed. As Achen and Snidal observed, "even if decision makers do not actually calculate, or if they rationalize their actions after the fact with foolish calculations, the assumptions of rational choice theory may yet remain true" and, therefore, useful. However, aside from Jervis, no one made an effort to evaluate, or even acknowledge, the potential contributions of deductive theory. This was particularly unfortunate because the debate about the superiority of one method over the other simply missed the point, since the question that should have been addressed was how both approaches could be enhanced to produce a more comprehensive body of evidence to test the theory. Similarly, Jervis criticized Achen and Snidal for concentrating on limitations of case-study research while offering few suggestions for improving the deductive theory of deterrence.⁴

One important implication of these persistent divisions is that the field inevitably becomes saturated with efforts to falsify theories by pointing out the flaws inherent in the methods used.⁵ Building on the *World Politics* symposium, the deterrence debate, with the rational choice and political psychology schools encompassing its two extremes, is presented in the next section as a case study of the problem. This particular debate was selected because it is multidimensional; it encompasses the divisions in the field between induction and deduction, single-case and multiple-case methodologies, and quantitative as opposed to qualitative approaches. As such, it provides a useful illustration of the impediments to progress that these divisions create and, more importantly, emphasizes the impact they have on the evolution of crisis-management theory.

CRISIS MANAGEMENT AND DETERRENCE: IMPEDIMENTS TO PROGRESS

The strength of rational choice theory and the decision models derived from its axiomatic base (game theory, expected utility models, deterrence models, and so on) has always depended on the degree to which the underlying assumptions of the theory offer at least a close approximation of reality. If decision makers, especially when faced with a foreign policy crisis, are incapable of processing information in a relatively straightforward manner, identifying alternatives and estimating their probability of success, assessing the impact of alternatives upon the values they seek to protect, and, in the end, selecting a course of action that promises to maximize expected payoffs (that is,

choosing the option they believe, at the time, to be in their best interest), then any theory derived from these assumptions would be trivial. This is precisely the point made by proponents of political psychology who have compiled what appears to be an impressive body of evidence against rational choice models. Included among the decision pathologies specified in the literature are perceptual limitations, poor information processing, and a limited search for alternatives or objectives (Holsti, North, and Brody 1968; Jervis 1976a; Janis and Mann 1977; Shlaim 1983; Stein 1985; Welch 1989); cognitive rigidity (Paige 1968; Stein and Tanter 1980; Lebow 1981); motivational biases (Lebow 1981; Jervis, Lebow, and Stein 1985; Steinberg 1989); "groupthink," or group conformity (Janis 1982, 1989; 't Hart 1994); and bureaucratic and organizational impediments (Allison 1971; Halperin and Kanter 1985; Art 1985; Sagan 1985). In sum, decision makers, particularly in a time of crisis, are either unwilling or unable to live up to the demands of rationality.

Of course, there is an abundance of conflicting evidence offered by proponents of rational choice that continues to fuel the debate, with each side highlighting cases that verify their version of reality while rejecting the findings and policy recommendations offered by the alternative school. In the absence of any attempt to identify areas of consensus, however, theoretical progress on the question of crisis management will likely remain elusive. The purpose of this study is to identify areas of consensus and to develop a more unified program of research.

Political Psychology vs Rational Choice

From the perspective of political psychology there are two principal decision models: the cognitive model and the motivational model. The cognitive model (Festinger 1957, 1964; Jervis 1976a; Holsti 1976; Lebow 1981) claims that rational decision making is often distorted through human cognitive limitations in information processing. More specifically, cognitive biases, a product of previously established belief structures, often predispose decision makers to accept and process information that conforms to expectations, while denying the significance of information that is inconsistent with those beliefs (Jervis 1976a). When confronted with new and contradictory information decision makers often reinterpret, change, deny, or even ignore it. Superficial alterations in a person's belief structure may take place, but major changes will occur only after enough evidence is compiled to clearly challenge preexisting beliefs (Jervis 1976a).

Whereas Jervis and others stress the ways in which cognitive processes distort decisions, other decision models emphasize the importance of "motivation" as a source of perceptual problems (Janis and Mann 1977; Lebow 1981; Lebow and Stein 1989a,b; Steinberg 1989). In many instances, within crisis situations, for example, decision makers are overrun with doubts and uncertainties and are generally reluctant to make irrevocable decisions. In order to deal with the overwhelming psychological stress associated with having to make these kinds of choices, leaders are prone to procrastinate, deny responsibility for the decision, or "bolster" (magnify the positive effects of the action regardless of its merits) (Janis and Mann 1977) – none of which is conducive to a rational decision process.

Despite their differences, both the cognitive and motivational models maintain that leaders are plagued by a number of decision pathologies, and a substantial body of empirical evidence has been generated to support these claims (Holsti and George 1975; Post 1991). Studies have found that policymakers in times of crisis rely on past experience to define present realities, engage in postdecisional rationalizations, remain insensitive to information that challenges previous commitments, use inappropriate analogies of the past for solutions to present problems, are often insensitive to warnings, misinterpret clues, fail to consider alternative points of view or seek out new information, are intolerant of ambiguity, tend to shift the burden of responsibility to the opponent, perceive their own options to be limited but their adversaries to be extensive, and so on.⁶

Two comprehensive case studies of decision making in World War I in 1914 and in the Cuban missile crisis of 1962 (Holsti 1965; Holsti, North, and Brody 1968), Jervis's (1976a) excellent research on misperception, and several studies of crises in the Middle East (Jervis, Lebow, and Stein 1985) have also contributed to the growing evidence against the rational actor model. Additional evidence from Janis's well-known work on groupthink (1982) has demonstrated that leaders often censor information and alternative viewpoints in order to preserve the emotional comfort of group unity. These pressures for group conformity are particularly strong during international crises when the use of military force is an option.

Critics of political psychology have identified several conceptual and methodological issues that should be assessed before making judgments about all of this evidence. A number of theories generated by this approach were borrowed in an ad hoc fashion from the field of psychology to explain important foreign policy decisions. Theorists have generally assumed that the situations under study in psychology and foreign policy were comparable, because in every case the decision-

making behaviour being explained was ultimately rooted in human psychology. But it is not clear that generalizations from one field to the other are appropriate or even realistic when one considers opportunities and constraints that apply in different settings. For instance, the original tests of stress and decision making were performed on subjects in controlled experiments, in simulations, or on businessmen in a corporate setting (Fleming 1991). Many of the results noted in Janis and Mann (1977) and Lebow (1981), two frequently cited works on the subject, were based on "reactions to emergency warnings about oncoming disasters that were matters of life or death, such as severe illness, radiation poisoning, earthquakes, tornadoes, floods, and air raids" (Janis and Mann 1977, 52). At best, each of these tests is only marginally analogous to interstate crises. Consequently, linking this research with, say, the behaviour expected from the United States and the USSR during intense international crises (for example, in Cuba in 1962) requires a more compelling defence.

Also, many of the original theories borrowed from psychology have since been discredited (or at least substantially qualified) by the same psychologists who performed the initial tests. Fleming (1991, 7) noted, for example, that research on groupthink actually "abated in the 1960s when experimental evidence failed to confirm the theories and when problems with the original experiments were discovered." Ironically, some of the earlier studies, rarely cited by proponents of the theory, demonstrated that conformity was positively related to low feelings of acceptance by the group.⁷ Although many of Janis' views have changed over time (see Janis 1989; Herek, Janis, and Huth 1987, 1989), his groupthink thesis continues to be widely cited and accepted as a legitimate criticism of rational choice theory. It is this widespread acceptance of a relatively weak thesis that should be questioned, especially considering Janis' slight change of heart in his later work (for example, Herek, Janis, and Huth 1987, 1989), which serves well to establish the point.

Controlled psychological experiments do not constitute the only source of empirical support for critics of rational choice theory. In-depth case studies are frequently put forward to illustrate the pathological nature of crisis decision making (Jervis, Lebow, and Stein 1985), although there are serious problems with this body of evidence as well: "Actors are rarely aware of the subtle psychological processes to which they are subject, and in cases where they are aware that such processes have contributed to a policy failure, they would be unlikely to admit to them" (Fleming 1991, 2). There is little direct proof, therefore, that such processes exist or have existed at all. More importantly, because many of the studies focus entirely on policy failures,

the evidence is often biased in favour of the propositions being tested. There is an unavoidable tendency, in other words, to select cases that prove the theory. At the same time, the alternative, rational choice model is usually operationalized, or defined, in a way that renders empirical verification almost impossible. Both Allison's (1971) and Stein's (1993, 32) depictions of the quintessential rational decision process, for example, stipulate a set of unrealistic expectations, especially in times of crisis. Contrary to assumptions held by these theorists, rationality does not require complete information or an exhaustive search for alternatives, objectives, and consequences; in fact, it is often irrational to undertake such a search if the marginal cost of continuing the investigation exceeds the marginal gain from doing so (Bueno de Mesquita and Lalman 1992). Similarly, rational choice theory does not stipulate that decision makers actually calculate the costs and benefits associated with each option, or measure the probabilities and utilities of every alternative outcome. This too would be irrational in most cases.

Another problem with nonrational models is the assumption that they somehow provide a stronger analytical tool for understanding and explaining international relations than models derived from the rational choice perspective (Verba 1969). It may be true that individuals (including world leaders) are influenced by a variety of subconscious psychological pressures – for example, cognitive rigidity (Jervis 1976a), a need for emotional security (Steinberg 1989), unconscious desires for affiliation and group conformity (Janis 1982), subjective motivational biases (Lebow 1981), a deep-rooted human drive to be aggressive (Freud 1938), or a compulsive, narcissistic, or paranoid personality (Post 1991). But it is very likely that the relative potency of these forces will vary from one individual to the next; presumably, some leaders are more aggressive, compulsive, or paranoid than others. Furthermore, different leaders will attempt to fulfil these needs in different ways, only one of which may be through public office or the state's foreign policy. Playing tennis five days a week, jogging three hours a day, firing staff, or even kicking the family dog are a few of the many different avenues available for venting frustrations and aggressions. For the sake of argument, assume that there is an equal probability that leaders and their foreign policy advisors fall into one of four categories described in Figure 1.1 (indirectly from Verba 1969). Although a very simplistic representation of reality, the example serves well to demonstrate the relative utility of rationality models in international relations and crisis management research. If foreign policy elites are more likely to rank high on the aggression scale and satisfy their need to vent their aggressive energy through foreign policy

11 Cumulation in International Relations

Foreign Policy Arena	Kick the Dog
High Aggressive 25%	High Aggressive 25%
Low Aggressive 25%	Low Aggressive 25%

Fig. 1.1 Hypothetical profiles of foreign policy elites

decisions, then nonrational models certainly would provide a very powerful analytical tool. On the other hand, if leaders are as likely to resemble any one of the remaining hypothetical profiles, then nonrational models would be inappropriate 75 percent of the time. In fact, there are several reasons why the percentage in the top left quadrant would be significantly lower, thus further limiting the utility of nonrational models. Experienced decision makers are likely to be more knowledgeable and skilful at handling the stress associated with having to make tough decisions. “The sort of progress which politicians have to make in any political system must make them used to high degrees of stress. If they are not able to act effectively under stress, they will not get very far in the profession ... Thus there is some form of indirect selection of effective crisis decision takers (Nicholson 1992, 129).” Also, to the extent that most crisis decisions are made by smaller groups uninhibited by bureaucratic or organizational influences, more, rather than fewer, alternatives are likely to be considered (Oneal 1988). In sum, foreign policy elites are subject to more of the conditions that inhibit the impact of personality and other subconscious psychological pressures.

Even some critics of rational choice have acknowledged its strength as a tool for analysing military-security crises. Time constraints during crises “effectively remove many of the informal channels of communication that operate in a non-crisis setting” (James 1993, 24). Also, leaders often receive support when other actors “rally around the flag” for efforts directed against an external threat (Levy 1989b). For these reasons, organizational and bureaucratic impediments are less potent.

Another problematic assumption of political psychology is that crises necessarily entail choices between alternatives that encompass complex value trade-offs (George 1985), which presumably makes crisis decision making much more difficult. But most crises involve only one principal “value” – to ensure that the situation does not escalate out of control or, if it does, to win the battle. Once again, domestic pressures that might produce conflicting objectives, frequently discussed in Lebow and Stein’s work on the Middle East

(1987, 1989a,b), may become much less relevant at the brink of war. In any event, the relative potency of these domestic forces depends on the political regime in question, because democratic leaders are more likely than military dictators to fall prey to these constraints. Even within democracies the impact of these domestic pressures will vary over time, usually in relation to electoral politics; the further they are from an election, the less influence negative polls and other societal constraints will have on crisis decisions. Finally, foreign policy elites, particularly in the West, have become much more effective at managing the flow of information through the media, and more confident of their ability to shape public opinion. In the United States, the Pentagon's control of coverage during the Gulf War is the most recent example (Hackett 1993; Mowlana 1993). It may no longer be essential to select a course of action that is suitable for domestic consumption, simply because domestic preferences can be moulded accordingly.

The most important criticism of the political psychology school, however, comes from a body of empirical evidence that paints a radically different picture of leaders in times of crises. The research agenda for a series of studies under the direction of Michael Brecher began with the following question: what is the impact of changing stress, derived from changes in perception of threats, time pressure, and the probability of war, on the processes and mechanisms through which leaders cope with crises and on their choices?⁸ The objective of the International Crisis Behaviour (ICB) project was to evaluate empirically many of the claims coming out of the Stanford studies on World War I (1914) and the Cuban missile crisis in 1962 (Holsti, North, and Brody 1968), and to offer an alternative image of decision making from that put forward by proponents of political psychology (see, for example, Jervis, Lebow, and Stein 1985; Lebow and Stein 1989a,b, 1990). Evidence from ICB case studies suggested that information processing and judgment actually improve during crisis situations. Each case offered a "resounding disconfirmation" of the "consensus" findings generated by the Stanford group and other psychological research (Richardson 1988): none of the ICB case studies showed an increase in cognitive rigidity; although leaders used lessons from the past as analogues to formulate policy, there was no evidence that this adversely affected the decision; in the Berlin blockade (1948) and two cases involving Israel (1967, 1973), the onset of crises actually led to an increased search for information and a more comprehensive evaluation of alternatives; there was no tendency to overlook options, and when fewer were considered it was due to time constraints or the availability of options at the point of decision.

Additional support for this model can be found in several other studies not connected with ICB research (for example, Snyder and Diesing 1977; Bueno de Mesquita 1981; Huth and Russett 1984, 1988; Herek, Janis, and Huth 1987, 1989; Bueno de Mesquita and Lalman 1992).⁹ In their evaluation of decision making in nineteen military-security crises since 1945, Herek, Janis, and Huth (1987, 1989) explored the relationship between seven decision-making pathologies and crisis outcomes. They found that in almost every case the quality of the decision (measured in terms of the favourableness of the outcome) was directly related to the vigilance displayed by officials during the crisis. More specifically, careful consideration of alternatives and objectives reduced the number of "avoidable" errors – for example, gross omissions in the evaluation of relevant information, of the costs and risks of preferred choices, of warnings, and so on. Their findings appeared to be consistent with expectations underlying rational choice theory.

Critics have noted several methodological and theoretical weaknesses with the Herek, Janis, and Huth study that should be addressed. In one particularly harsh critique, Welch (1989) claimed that the conclusions derived from analysis of the nineteen crises were biased. In the case of the Cuban missile crisis, for example, five of the symptoms (or decision pathologies) that were coded by the authors as being absent were actually present during the executive committee meetings. According to Welch, the president and his advisors failed to discuss or even consider viable alternatives and objectives: "Despite the range of alternatives tabled for discussion, only two received sustained scrutiny by the president and his advisors: the limited quarantine and the massive air strikes. All diplomatic options were quickly dismissed because the president's advisors could not imagine a non-military first move" (1989, 432). Similarly, there was really only one primary objective discussed by the executive committee: "to secure the withdrawal of Soviet missiles from Cuba." Welch maintained that "there is no evidence ... of debate on the necessity of this objective. Nor is there any evidence of sustained consideration of why Soviet nuclear missiles in Cuba were unacceptable" (433). The important theoretical lesson from Welch's study is that defects in the decision process did not appear to have an appreciable affect on the outcome, which according to most observers was quite favourable. Process, in other words, is not necessarily related to performance.

In response to Welch's allegations, Herek, Janis, and Huth (1989) published a persuasive defence of their original coding decisions. In each case, they offered compelling evidence that Welch either misinterpreted their arguments or misrepresented their methodology.

According to them, Welch outlined criteria that they did not use and then proceeded to describe how Kennedy and members of the executive committee failed to live up to those standards. With respect to Welch's so-called new evidence, Herek, Janis, and Huth argued that it did not "substantially contradict (their) ratings of the quality of U.S. decision making during the Cuban missile crisis," nor was it sufficient to alter their overall conclusions regarding process and outcome. The evidence presented in their response was both solid and convincing.

Unfortunately, notwithstanding cases in which decision makers have correctly interpreted their opponents' signals, have been conscious of the potential for decision pathologies, and have ultimately chosen the appropriate course of action (as in the Cuban missile crisis), the image of a room filled with wide-eyed, trembling, coffee-drinking smokers remains a common one in much of the literature. But the assertion that decision makers are very likely to fall prey to subconscious pressures and less likely to decide rationally requires a more compelling defence. The mere likelihood that decision pathologies will take over during a crisis is not sufficient to render rational choice models trivial. Proponents of political psychology rarely explain why their models provide a more plausible set of assumptions upon which to develop crisis management recommendations and foreign policy proposals.

One tentative conclusion (perhaps a synthesis of arguments) coming out of the ongoing debate between proponents of rational choice and political psychology is that stress actually may have a positive impact on the decision process, but only up to a point. In other words, there may be a curvilinear relationship between stress and performance. Even critics of rational choice accept the potential explanatory power of the inverted-U hypothesis (George 1985, 507) displayed in Figure 1.2. This hypothesis would certainly go a long way towards explaining the apparent contradiction in the findings. Unfortunately, instead of serving as a ground for consensus and synthesis, the conflicting evidence continues to fuel the debate, with each side reemphasizing research that verifies their version of reality. In fact, the debate has simply shifted to the precise nature of the inverted U: How steep is the incline and decline of the curve? Where exactly is the top of the curve in relation to the x- and y-axes? Where does the decision maker reside on the curve, and under what circumstances?¹⁰

Implications for Developing Crisis Management Theory

Given the lack of consensus on the relationship between crisis-induced stress and performance, theoretical progress in the area of crisis

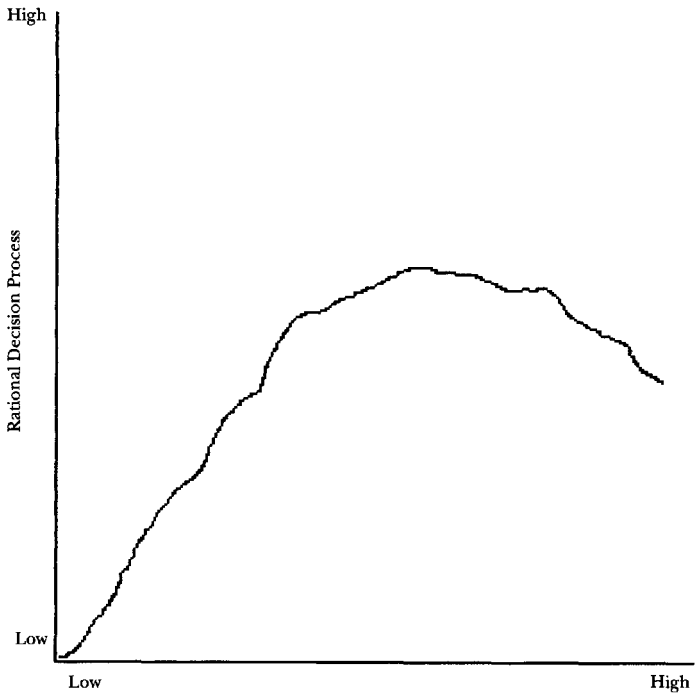


Fig. 1.2 Rationality and stress in foreign policy crises

management will remain elusive. After all, crisis management theory ultimately rests on assumptions about the decision-making process. Depending on which set of assumptions and body of evidence one accepts as valid, one's descriptions, explanations, and predictions of political behaviour will vary. More importantly, policy recommendations about appropriate management strategies will differ. If one accepts the assumptions underlying rational choice theory, for example, then deterrence – a strategy derived largely from rationality assumptions – is an ideal crisis management technique. The study of deterrence theory, then, becomes a legitimate and integral part of the research agenda of the field, with the central question being how best to achieve or avoid a deterrence success or failure. According to conventional wisdom, commitments must be clearly defined and communicated to the opponent, one's resolve to defend those commitments must be credible, the deterrer's capabilities to follow through on the counterthreats must be unquestioned, and so on (Brodie 1959a, b; Schelling 1960, 1966).

On the other hand, if one accepts the evidence from political psychology, then deterrence strategies may not be effective at all. In many cases, coercive tactics like deterrence may actually provoke hostile reactions from an opponent faced with domestic pressures (Lebow 1981; Lebow and Stein 1987). Studies by Lebow and Stein indicate that even in cases where deterrence was executed correctly (that is the commitment was communicated, the credibility of the counterthreat was clear, resolve was unquestioned), the retaliatory threat failed, and the opponent attacked (for example, the Yom Kippur War, 1973; the Falkland Islands War, 1982). Consequently, not only does deterrence fail as an effective crisis management technique, but it also fails as a theory of crisis behaviour. As Kolodziej points out, "Leaders may actually choose to defeat an adversary's best deterrent efforts ... Conversely, policies pursued in the name of deterrence ... may, at best, be benign, having no appreciable effect on an adversary ... Worse, a deterrer's moves may precipitate a military buildup by an opponent or prompt an armed clash" (1987, 124-5).

Criticising deterrence theory is a central feature of the Lebow-Stein research program.¹¹ Their attempt to identify the problems that plague the central assumptions of the theory, the methodological errors embedded within a great deal of deterrence literature, and the practical problems associated with implementing deterrence as a strategy constitute the main thrust of their work.¹² "Rational deterrence theories (pertain to) nonexistent decision makers operating in nonexistent environments" (Lebow and Stein 1989a, 224).

The most important policy recommendation coming out of this body of work is that leaders should avoid coercive strategies that entail explicit deterrent threats, precisely because they are likely to provoke the nonrational behaviour officials are attempting to deter. Instead, emphasis should be placed on correcting many of the decision pathologies noted earlier: leaders should refrain from accepting simplistic images of the adversary, ignore stereotyped interpretations of historical events, disregard typical attacker/defender, good side/bad side designations, attempt to view the crisis from the opponent's perspective, and so on. Reassurance, not deterrence, is the preferred policy course. This strategy entails a more cooperative approach to crisis management that recommends conciliation in an effort to create an environment of mutual trust and reduce the chances of misperception and escalation.

There is an important inconsistency in the argument for a reassurance strategy that should be noted, however. Presumably, a rational decision-making process is a prerequisite for either strategy to be

effective. Both reassurance and deterrence require that each side in a dispute accurately interpret the adversary's actions and respond accordingly. But there is no logical reason why a leader at the brink of hostilities would be less likely to misperceive or misinterpret a cooperative act. With a long history of mistrust, for example, conflicting parties will often see "manipulative intent" in actions that appear to be conciliatory (Lindsfold, Betz, and Walters 1986). The penchant for mistakes and misperceptions, in other words, is not a function of an opponent's actions but of the history of a relationship and the stress the leaders experience during the crisis. Ironically, Lebow and Stein are therefore forced to defend the rationality assumption in order to emphasize the benefits of reassurance as a crisis management technique. In addition, reassurance strategies are as likely to be disastrous when used under the wrong circumstances, and domestic pressures that occasionally hamper a well-formulated deterrence policy are just as likely to jeopardize a reassurance strategy.

If the rational choice and political psychology schools are each correct about some things, and if there is a curvilinear relationship between stress and the decision process (fig. 1.2), then what we need to accommodate both schools is an integrated decision model that incorporates both sets of assumptions and bodies of evidence. Unfortunately, the ongoing debate has done more to restrict theory by inhibiting our ability, as a community of scholars, to acknowledge progress when it occurs and by creating obstacles to developing a more progressive research agenda that builds on previous work through identification of consensus. With this in mind, the following section will attempt to contribute to the search for cumulative knowledge about crisis management by highlighting areas of consensus in the field that are often overlooked or ignored by both sides in the debate about rationality.

SHIFTING EMPHASIS: AN ALTERNATIVE APPROACH TO CUMULATION

The purpose of the preceding section was to demonstrate how a division between rational choice and political psychology, each favouring a particular approach to the study of crisis behaviour while rejecting the utility of methods used by the counterpart, has affected the cumulation of crisis management theory in two ways: first, by inhibiting our ability, as a community of scholars, to acknowledge progress when it occurs; and second, by creating obstacles to developing a more progressive research program that builds on previous work.

The fundamental error implicit in these debates is the assertion that contributions to knowledge can be evaluated solely on the basis of the approach that is used. Contrary to this assertion, this study assumes we should classify, not accept or reject, research on the basis of methodology, and evaluate contributions within the confines of a specific methodological program. This approach is expected to avoid counterproductive debates over methods and facilitate cumulation by identifying areas of consensus often overlooked or ignored by critics, and it serves as the starting point for chapter 2. Consistent with this organizing strategy, U.S.-Soviet rivalry is isolated for more in-depth evaluation, with nuclear deterrence as the main focus of inquiry. The central objective is to synthesize what we have discovered about the conduct of these powers during intense crises and, on the basis of the investigation, suggest new avenues for aggregate research that build on previous work, the focus of chapters 4-7.¹³

2 Nuclear Deterrence: The Record of Aggregate Testing

Progress in the debate over rational deterrence theory, both conventional and nuclear variants, has always depended on the ability of scholars to identify a body of evidence that would be appropriate for testing a wide range of propositions derived from the theory. Unfortunately, notwithstanding the tremendous amount of time and energy spent on producing a suitable list of cases and notwithstanding several noteworthy surveys of the literature (Jervis 1979; Jervis, Lebow, and Stein 1985; Levy 1988; Pages 1991; Huth, Gelpi, and Bennett 1993; Harvey and James 1992; Harvey 1995), cumulative knowledge about deterrence, both as a theory and as a strategy, remains elusive. It still is unclear whether decision makers have acted according to the logic derived from standard applications of the theory. Moreover, the most prominent aggregate testing strategy, originally designed by Huth and Russett (1984, 1988, 1990) and later criticised and revised by Lebow and Stein (1987, 1989a,b, 1990), continues to be plagued by ongoing debates over methods and case listings. Lingering divisions over the coding of deterrence successes and failures have become counterproductive, primarily because each side can produce evidence to support its own interpretation of events (Harvey 1995). Although debates over the accuracy of historical accounts have been constructive, very little effort has been directed towards developing alternative testing strategies that lie outside the success/failure framework or looking at a wider range of propositions derived from the theory.

This chapter unfolds in four stages. First, major problems common to the aggregate research program on nuclear and conventional deterrence are described. Next, criteria of reliability and validity are used

to evaluate contributions and shortcomings of the dominant testing strategy. Then, a case study of the Syrian invasion of Lebanon in 1976 and Israel's threat of military retaliation highlights the many obstacles associated with coding any deterrence encounter and, more importantly, illustrates how, in the context of the dominant testing strategy, the same crisis can be used both to support and refute the theory. The final section confirms the importance of identifying the universe of cases appropriate for evaluating deterrence, but recommends a different strategy that employs a less rigid, but no less valid, set of selection criteria.¹

THE CONTEXT OF AGGREGATE ANALYSIS

One straightforward approach to the classification of deterrence literature is that of Smith (1981). The two-dimensional scheme presented by Smith is sufficient to distinguish the subset of studies that is most relevant. Table 2.1 displays the taxonomy, with representative examples. The first dimension focuses on the degree of structure exhibited by the analysis. More specifically, as explained by Jervis (1979, 289), deterrence literature can be divided into three "waves," ranging from investigations that deal only with the basic concepts and underlying logic of the theory (first wave), through those that advance some generalizations and historical applications (second wave) and, finally, to studies that confront these generalizations with empirical evidence (third wave) (Smith 1981, 26). As for the second dimension, deterrence literature can also be categorized according to the number of states being examined. This produces "a continuum which ranges from zero for studies which are completely abstract to an arbitrarily large sample representing the maximum number of nations present in the system at any point" (25).

Based on degree of structure and empirical range, the studies most relevant to this review appear in the final cell of Table 2.1.² It is interesting to observe that the fifteen studies listed as third wave, with a multistate sample of cases, include all of those that could be located. By contrast, each of the other eleven cells in the matrix contains one or more representative instances. Among the fifteen studies just noted, seven have developed a list of cases deemed relevant to nuclear and conventional deterrence theory: Russett (1963), George and Smoke (1974), Organski and Kugler (1980), Kugler (1984), Huth and Russett (1984), Huth (1988a) and Betts (1987). Fink's (1965) study stands as a replication of Russett (1963), while Huth (1988b, 1990) and Huth and Russett (1988, 1993) rely on the data set presented in Huth (1988a).³ The remaining four studies (Weede 1981, 1983;

Table 2.1
Classification of Deterrence Literature

<i>Degree of Structure</i>	<i>Empirical Range</i>			
	<i>Abstract</i>	<i>Specific to One State</i>	<i>Specific to a Dyad</i>	<i>Multistate Samples of Cases</i>
First wave	Brodie (1946, 1959a,b) Schelling (1960, 1966) Kenny (1985)	Garthoff (1962)	Milburn (1959) Kahn (1961, 1962) Walters (1974) Freedman (1986) George (1986)	Brodie (1973)
Second wave	Kaplan (1958) Snyder (1960, 1971) Ellsberg (1961) Deutsch (1968) Brams (1975, 1985) Gauthier (1984)	Jabber (1971) Subrahmanyam (1974)	Wohlstetter (1959) Snyder (1961) Maxwell (1968) Sokolovsky (1975) Jervis (1980) Wagner (1982)	Jervis (1976a,b)
Third wave	Morgan (1977) Nicholson (1992) Powell (1990)	Wohlstetter (1962) Russett (1967) Paige (1968) Blechman, Kaplan, et al. (1978)	Doran (1973) Brams and Kilgour (1987a,b)	Russett (1963) Fink (1965) George and Smoke (1974) Organski and Kugler (1980) Weede (1981, 1983) Bueno de Mesquita and Riker (1982) Kugler (1984) Huth (1988a,b, 1990) Huth and Russett (1984, 1988) Betts (1987) Geller (1990)

Source: Adapted from Smith (1981, 23).

Bueno de Mesquita and Riker 1982; Geller 1990) tested propositions through different approaches. Weede (1981), for example, compared the relative frequency of war in dyads where extended deterrence did and did not apply. He assumed that in all cases extended deterrence would pertain to superpowers and their allies. Although Weede's alternative testing strategy did not isolate individual cases of deterrence, studies such as his will also be discussed.

Before evaluating the seven studies that have produced case lists, some boundaries will be set for the analysis. First, it should be recalled that an extensive body of literature has dealt with deterrence by focusing on individual cases.⁴ Although such studies provide a great deal of insight into the process of decision making in crises, they employ a wide range of models and "prove [to be] little more than suggestions for theory building" (George and Smoke 1974, 93). Consequently, individual cases will not be treated as data sets in this review. Second, while mathematical models have been used to derive and test propositions embedded within deterrence theory, these deductive entities consist primarily of abstract analysis of expected behaviour and provide little systematic evidence regarding actual conduct in deterrence cases.⁵

Neither of the two approaches noted above – mathematical models and intensive case studies – has provided the kind of verification associated with quantitative analysis of a large number of cases (George and Smoke 1974). While it is not the intention here to debate which method is more appropriate for developing theory, it is apparent that large-N research has been relatively uncommon. The primary concern of the following review, then, is with the contributions made by this nascent program of aggregate study. Specifically, emphasis will be placed on assessing nuclear deterrence and identifying a more integrated approach towards testing. Multiple-case compilations that focus strictly on conventional deterrence, therefore, will not be reviewed.⁶

NUCLEAR DETERRENCE AND AGGREGATE DATA

Russett's (1963, 98) analysis of seventeen crises between 1935 and 1961 appears to be the first study of deterrence based on multiple cases. Each crisis in his data set conformed to the necessary and sufficient conditions for a situation of extended deterrence. Russett included "all the cases during the last three decades where a major power 'attacker' overtly threatened a pawn with military force and where the defender either had given, prior to the crisis, some indication

of an intent to protect the client or made a commitment in time to prevent the threatened attack." In other words, deterrence was extended by the defender to the pawn. According to Russett, deterrence succeeds when "an attack on the pawn is prevented or repulsed without conflict between the attacking forces and the regular combat units of the major power (defender)." If the attacker finds fault with either the intensity or credibility of the defender's threat, an attack ensues and deterrence has failed. Among Russett's seventeen cases, six are coded as successes and the remaining eleven as failures.

Russett (1963, 103) assessed the impact of several potential determinants of credibility, including the "economic, political, and military interdependence of pawn and defender." If a given variable, such as economic ties between defender and protégé, past or present military integration, local military capabilities, or something else, was present in an overwhelming majority of the cases in which deterrence succeeded, and absent when deterrence failed, it was considered an important determinant of credibility. Among the independent variables, economic ties had the most observable connection with successful deterrence.

About one-half of Russett's cases involved a defender that had previously detonated a nuclear device. It should be noted that six of the nine cases involving a nuclear defender resulted in successful deterrence. However, because Russett did not examine the potential impact of nuclear weapons on the attacker's decision, the relationship between those forces and successful deterrence cannot be inferred directly.

Finally, the categories of the dependent variable – success or failure – reflect whether or not war (or other major hostile interactions) has broken out. Effective deterrence, however, also implies that the defender has achieved at least some of its political and military objectives in the dispute (Organski and Kugler 1980; Kugler 1984) or has successfully prevented the aggressor from attaining its policy goals (Huth and Russett 1984). Since Russett focused exclusively on violence in defining the outcome variable, the generalizations are less relevant to the issue of goal achievement. A more detailed explanation for the inclusion of both criteria is presented in chapter 3, which looks at new directions for aggregate research.

In an effort to develop a typology of deterrence failure, George and Smoke (1974, 537–44) identified eleven major cases of deterrence failure by the United States between 1948 and 1962: (1) the Berlin blockade, 1948; (2) the outbreak of the Korean War, 1950; (3) the Chinese Communists in Korea, 1950; (4) Korea/Indochina, 1953–54; (5) Taiwan, 1954–55; (6) the Hungarian Revolution, 1956; (7) the

Middle East, 1957–58; (8) the Quemoy crisis, 1958; (9) the Berlin deadline crisis, 1958–59; (10) the Berlin *aide-mémoire* crisis, 1961; and (11) the Cuban missile crisis, 1962. They found at least one of three types of deterrence failure to be present in each instance: *fait accompli*, limited probe, and controlled pressure. When the initiator believes that the defender has made no commitment with respect to the area in dispute, the *fait accompli* failure occurs. This belief results in a maximum effort by the initiator to “achieve the objective quickly so as to deprive the defender of the time and opportunity necessary to reverse his policy of no commitment.” The limited probe failure occurs when the initiator believes that the defender’s commitment is “uncertain.” In such cases, the option chosen by the initiator is a controlled application of limited force that requires a clarification of the defender’s interests. In a case of controlled pressure, the third type of deterrence failure, the initiator believes the defender’s commitment is “unequivocal” but soft. Thus the initiator chooses to apply pressure to convince the defender that it would have “great difficulty and incur unacceptable risks” if it attempted to honour the commitment (George and Smoke 1974, 537, 541–4).⁷

Since at least one of the three types of deterrence failure occurred in each of the cases identified, it can be inferred that crisis policy-making, within both the defending and the potential initiating power, is a process composed of a series of interrelated decisions. Hence, deterrence is likely to succeed or fail through time and stages. Given the complexity of the situations, there is likely to be more than one causal pattern through which deterrent threats may work. In addition, the findings of George and Smoke show that nuclear weapons have little to do with either the effectiveness or the credibility of the retaliatory threat; deterrence has often failed regardless of the strategic nuclear capabilities of the parties involved. In fact, judging from the set of cases identified by George and Smoke, the two central determinants of success or failure appear to be the extent to which the defender is committed to the area in dispute and the subsequent perception of that commitment on the part of the attacker.

There are some complicating factors worth noting in George and Smoke’s research design. Most significantly, they neither included guidelines regarding the selection of evidence nor specified whether their cases dealt with “immediate” or “general” deterrence. In a setting of general deterrence “opponents ... maintain armed forces to regulate their relationship even though neither [side] is anywhere near mounting an attack”; in immediate deterrence “at least one side is seriously considering an attack while the other is mounting a threat of retaliation in order to prevent it” (Morgan 1977, 28). Moreover,

there are two further types of strategies to consider: deterrence directed towards preventing an attack on oneself or an attack on another party – an ally, client state, or friendly neutral (Huth and Russett 1984, 496). The latter strategy is usually referred to as “extended” and can be separated still further into “extended general” and “extended immediate” deterrence.

Most observers agree that the success or failure of general deterrence is extremely difficult to evaluate. As Huth and Russett (1984, 497) point out, the absence of an assault does not necessarily attest to the success of general deterrence. The potential attacker may never have intended to attack in the first place, or the priorities of the potential attacker may have changed for other reasons.

Given these difficulties, it is clear that George and Smoke should have isolated a common type of deterrent situation; one should not necessarily mix apples and oranges on the assumption that the fact that they are both examples of fruit is what matters the most. Some examples from George and Smoke’s collection of cases should help to bring out the different interpretations that depend on whether general or immediate deterrence is at issue. In the Hungarian crisis of 1956, for example, the United States actively pursued a strategy of containment but never explicitly threatened retaliation in response to the Soviet invasion plans. Although the general deterrent, in the form of the U.S. containment policy, may have failed, it is questionable whether the crisis ever reached the stage of immediate deterrence. (Even the failure of general deterrence is, in fact, debatable; after all, the USSR did not expand its sphere of influence.)⁸ It would therefore be inappropriate not only to include Hungary among cases of immediate deterrence but also to use it for the purpose of developing a typology of immediate deterrence failure.

Similarly, the Cuban missile crisis provides an example of a general deterrence failure, an immediate deterrence success, and an immediate “compellent” success. Whereas deterrence is used to prevent the occurrence of undesired actions (that is, to maintain the status quo) compellent threats are used to force opponents to do something they would not otherwise have done. General deterrence had failed once the USSR placed missiles on Cuban soil. However, the United States discouraged the USSR from completing the deployment by threatening retaliation if the Soviets continued. Indeed, the implementation of the American blockade resulted not only in a halt to missile deployments but ultimately in the withdrawal of Soviet missiles from the island (an apparent compellent success). As in the case of Hungary, it is questionable whether Cuba should have been included in the list of deterrence failures compiled by George and Smoke.

Without an explicit definition of the intervening variable – type of deterrence or compellence – any generalizations made about success or failure are problematic.

Organski and Kugler (1980) identified fourteen cases of potential nuclear deterrence between 1945 and 1979. They used this data set to test whether crises have been resolved according to the logic derived from standard nuclear deterrence theory. If so, according to Organski and Kugler, states with nuclear capabilities should have achieved their goals in an overwhelming number of cases. Most notably, the nuclear power “should have [had] its way in the central issue in the dispute” (163). Among other things, nuclear deterrence theory entails greater success for the antagonist that possesses qualitatively superior capabilities. This expectation, however, was not fulfilled by the data. Among the fourteen cases, a nuclear power confronted a nonnuclear adversary in seven – the most direct test of nuclear deterrence – and in only one such instance (the Berlin blockade) was the nuclear power considered to have won (176).

One shortcoming of Organski and Kugler’s study is its mixture of cases of immediate deterrence with cases of general deterrence. Like the investigation by George and Smoke, this treatment does not specify the “universe” from which the cases have been selected. Organski and Kugler did not define precisely the type of deterrence they were considering; nor did they specify whether a nuclear threat occurred in any of the fourteen crises. Consequently, identification of the factors contributing to the success or failure of deterrence becomes problematic. For future reference, it is important to be aware of the level of coercion threatened by a defender, because that could play a decisive role in shaping the outcome.

Kugler (1984, 477), in a follow-up study, identified fourteen cases of potential nuclear deterrence between 1948 and 1979. “To assure that the threat of nuclear weapons played a direct role in each confrontation,” Kugler included “only crises in which nuclear weapons are judged to have played a central role.” Initially, the set of cases had incorporated forty-two crises with potential for escalation into a nuclear confrontation, but Kugler used the following criteria to reduce the collection. Extreme crises were those judged by Butterworth (1976, 484) as ones that might possibly lead or ones that were likely to lead major powers to engage in a war using nuclear weapons, and that were isolated by the Center for Arms Control and International Security (CACI) as potentially or actually severe because they achieved the highest scores on questions concerning three interlocking variables: the nuclear capability of contestants, the degree of involvement in the crisis, and the threats made as the conflict developed.

Implementation of these criteria resulted in some of the crises from Organski and Kugler's (1980) study being replaced; the findings, however, remained the same. Kugler (1984, 470) found that nuclear powers obtained policy objectives in only five of the fourteen crises (35.7 percent). He concluded that "nuclear deterrence may be a useful political strategy, but it may not provide an effective empirical description of the behaviour of nations faced with the prospect of major war." However, it is entirely possible that the success rate for cases not meeting criteria (a) and (b) is much lower. If so, that would suggest that a strategy of deterrence is, in general, very difficult to pursue and that the results do not reflect a specific failing of nuclear powers in crisis management.

Huth and Russett (1984) expanded Russett's original sample of seventeen to include fifty-four cases of extended deterrence.⁹ In this expanded version of Russett's original study, additional variables – including the defender's strategic nuclear forces – were incorporated. Based on these presumed determinants of expected value, Huth and Russett (1984, 509) derived a set of hypotheses. Tapping aspects of an expected utility model, these hypotheses span three general categories: relative power, past behaviour in signalling current intentions, and the contextual importance of communitarian ties between states.

Evaluation of the outcome of a case focused on whether a specific attack occurred, whether the attacker gained its principal goals with minimal fatalities, and whether the attacker occupied the protégé's territory for several years. Results from the updated data set resembled those of the original study. Huth and Russett found that, of twelve hypotheses, the evidence supported three: An attacker will be

- more likely to fight to the degree that the attacker's existing local military capabilities exceed those of the defender,
- less likely to fight the stronger the economic linkages between defender and protégé,¹⁰
- less likely to fight the stronger the political-military linkages (arms transfers) between defender and the protégé.

It is interesting that the following subset of hypotheses turned out to be inconsistent with the data. An attacker will be

- more likely to fight to the degree that the attacker's overall strategic-military capabilities exceed those of the defender,
- less likely to fight if the defender is known to possess nuclear weapons,
- more or less likely to fight if the defender has fought in the past (both hypotheses were tested).

As Huth and Russett (1984, 524) pointed out, the two sets of findings suggest that "a definition of deterrence as primarily sensitive to strict calculations of military capabilities is both mistaken and profoundly dangerous." Thus it is not unreasonable to explore the roles played by factors other than strategic nuclear superiority in making or breaking deterrence.

Compared to Russett's original study, Huth and Russett (1984) provided more comprehensive treatments of the independent and, especially, dependent variables. Success, for example, no longer depended solely on whether war (or major hostile acts) had occurred but also included the attacker's achievement of political and military objectives. In addition, the fifty-four cases of immediate deterrence in the data set satisfied four clearly specified necessary and sufficient conditions, thereby contributing to the probable reliability of the findings.

Notwithstanding these improvements, several matters related to testing require further attention. For instance, within the set of propositions, relative potency is not assigned explicitly to the explanatory variables. What happens, for example, if the local military balance favours the attacker, but the economic or military ties between defender and protégé are substantial? Which of these variables is considered more important when calculating expected utility for the attacker? Furthermore, what part, if any, do the less significant variables play in the attacker's decision-making process? If such factors are potentially causal, then an assessment of relative importance is required.

Huth and Russett assumed that a credible threat existed in all cases of successful deterrence. It therefore could be inferred that an incredible threat preceded every failure. However, as Fink (1965) pointed out, that line of reasoning is questionable.¹¹ In fact, deterrence can fail regardless of the credibility of the defender's threat. A perfectly credible threat of retaliation may be completely ineffective if the costs to the attacker of inaction are so high that they overwhelm the costs of using force. For example, although the Japanese, prior to attacking Pearl Harbour, "perceived the U.S. threat to be highly credible," they attacked anyway, further demonstrating the fallacy of equating credibility with effectiveness (Levy 1988, 500).

A final problem, anticipated by Russett (1967, 94), is that two substantial inferential leaps are implied in the conclusion that the bonds between defender and protégé determine the credibility of the retaliatory threat. One leap is from the existence of such links to their perception by the attacker, and the second is "from their perception to their entry in a significant way into the [attacker's] calculations."

As Russett observed, "we are precisely in the middle of the old problem of confounding correlation with causation; the mere presence of the bonds does not prove that they influenced the outcome." Nevertheless, all other things being equal (such as the costs facing the potential threatening power as a result of continuing inaction), greater credibility should be associated with a higher probability of successful deterrence.

Huth (1988a,b) analyzed fifty-eight cases of extended immediate deterrence. His findings were similar to those of Huth and Russett (1984). Specifically, the immediate and short-term balance of forces, as opposed to the long-term correspondence, affected deterrence outcomes. Furthermore, possession of nuclear weapons by the defender "did not have a significant impact on deterrence outcomes when the target was a non-nuclear power" (Huth 1988b, 435). One noteworthy difference from the 1984 study did emerge: high levels of economic and political-military ties between defender and protégé could not be linked to deterrence success.

In a related study, Huth and Russett (1988) analyzed Huth's fifty-eight cases of immediate deterrence using a somewhat different set of independent variables. One overall conclusion, central to the current study, did not change: nuclear weapons could not be linked to the success or failure of deterrence.¹²

Focusing more specifically on the role of nuclear weapons in the deterrence process, Betts (1987, 213) attempted to answer three important questions: How often and why did leaders resort to nuclear threats? How useful were those threats in affecting the outcomes of crises? and How did the nuclear balance of power figure either in decisions to threaten or in the results? For data he relied upon East-West crises "in which some sort of nuclear threat was made by one or both sides." Betts included crises at the "margins" of the basic super-power conflict, referring to "those gray areas that were neither clear, central interests of one side that it would be expected to defend to the death, nor ones so obviously peripheral that no one would expect military engagement after them" (16). In order to assess the relationship of the nuclear balance of power to the issuing of threats, Betts grouped the cases "according to the degree of nuclear risk facing the country making the threat—that is, how much vulnerability to nuclear retaliation it faced if its first-use threat had to be executed" (17). The cases, which span the years 1948–80, are divided as follows: In the low-risk category are (1) the Berlin blockade, 1948; (2) the Korean War, 1950–53; (3) Asian crises, 1954–55; (4) Suez, 1956; (5) Lebanon and the Taiwan Straits, 1958; and (6) the Soviet-Chinese border clashes, 1969. In the high-risk category are (7) the Berlin deadline crisis,

1958–59; (8) the Berlin *aide-mémoire* crisis, 1961; (9) the Cuban missile crisis, 1962; (10) the Middle East War, 1973; and (11) the Carter Doctrine, 1980. Betts' central findings (213–14) can be summarized as follows:

- 1 Between 1945 and 1965, U.S. presidents “were inclined to introduce vague nuclear threats in military confrontations.”
- 2 U.S. nuclear superiority did not instil confidence in political decision makers.
- 3 Nuclear parity did not significantly change the behaviour of leaders; “it fostered anxiety about whether first-use threats remained practical but did not prevent resurrection of the tactic when crises comparable to former cases occurred.”
- 4 With respect to crises, the nuclear balance “appears to have played a moderately influential role when it was uneven and an uncertain one when it was equal.”

These generalizations do not have uniform implications regarding the effectiveness of nuclear weapons. The role played by those armaments seems to depend very much on the situation at hand.

Using the 1988 data set (with marginal adjustments), Huth (1990, 271, 276–7) examined the deterrent role of nuclear weapons in “international confrontations not characterized by MAD” (mutual assured destruction). He found that, when the defender in a case possessed “the means to deliver a nuclear strike against the potential attacker,” a positive but statistically insignificant connection emerged with successful deterrence. Thus Huth concluded that nuclear weapons “do not seem to play an important extended deterrent role in the historical cases analyzed.”

Aside from the multiple-case studies just reviewed, there are some alternative approaches toward testing nuclear deterrence theory on an aggregate basis. Bueno de Mesquita and Riker (1982), Weede (1981, 1983) and Geller (1990) have focused on the more general question of whether nuclear deterrence has reduced the risk of war. It is important to note that these studies are not concerned with the specifics of what makes deterrence work. Instead, the deterrent value of threats of military retaliation is at issue, a question posed by Raser (1966, 314) but seldom addressed. Although the focus of this final subset of investigations is different, the overall objective remains much the same: to assess whether the basic assumptions of the theory are consistent with the behaviour of states in the real world. Efforts toward that end have examined the extent to which nuclear threats appear to have prevented the onset of war and other major hostilities. While

these data-based studies do not isolate particular cases, some important insights are derived about deterrence. In fact, contrary to the thrust of the findings presented earlier, the majority opinion here is that nuclear deterrence theory is supported by the evidence; wars and major hostilities seem to have been prevented as a result of nuclear threats.

Bueno de Mesquita and Riker (1982, 283, 289), to begin, observed that "in a conflict between a nation with a nuclear capability and one without, the conflict may become nuclear, while, in a conflict between two nuclear powers, the conflict is unlikely to become nuclear, by reason of fear of retaliation."¹³ They used indicators provided by Gochman (1975) and Singer and Small (1972) to measure the three types of conflict that could develop: threat, intervention, and war. Threats involved "mere verbal assaults, without escalation to violence." Interventions corresponded to "those conflicts in which modest levels of violence were used," while wars included conflicts "in which at least one thousand battle-related fatalities occurred." As Bueno de Mesquita and Riker pointed out, "the logic of deterrence" suggests that as the level of nuclear constraint decreases, the resulting conflict type will be more intense, as revealed by Table 2.2. This, they observed, is true "because each row represents a successive decrease of a deterrent threat that is expected to restrain conflict behaviour." More specifically, the evidence shows that no wars occurred between states with conflicts falling in the first two rows of Table 2.2. These states experienced a "disproportionately large percentage of nonviolent, verbal conflicts ($p < .01$)"; by contrast, those engaged in conflicts that did not involve a nuclear restraint "experienced a disproportionately high probability" of war and a "disproportionately low probability of finding their conflicts confined to verbal threats (both p -values $< .01$)" (1982, 290, 291). Like others before, this study focused strictly on the degree of violence experienced by the adversaries to assess the outcome.

Weede (1981, 1983), in an effort to find out "whether mutual nuclear deterrence and its extension to superpower allies or clients works," examined 299 politically and militarily relevant dyads between 1962 and 1980. According to Weede (1983, 237), if extended deterrence works, then it should be observable among relevant pairs, that is, among "the subset of dyads where the risk (possibility) of war is substantial." He compared the relative frequency of war in dyads with and without the potential for extended nuclear deterrence, respectively. ("War" followed the definition provided by Singer and Small (1972); see also Small and Singer (1982).) According to Weede (1983, 237), deterrence applied "between the U.S. and its allies on the one hand and the USSR and her allies on the other hand." In the

Table 2.2
Nuclear Constraints and Conflict Type

<i>Nuclear Constraints</i>	<i>Conflict Type</i>		
	<i>Threats</i>	<i>Interventions</i>	<i>Wars</i>
Nuclear power vs nuclear power	4 (0.67)	2 (0.33)	0 (0.0)
Nuclear power vs nation with nuclear ally	7 (0.54)	6 (0.46)	0 (0.0)
Nuclear power vs non-nuclear power	8 (0.35)	13 (0.57)	2 (0.09)
Non-nuclear power vs non-nuclear power	10 (0.17)	31 (0.53)	17 (0.29)

Source: Bueno de Mesquita and Riker (1982, 290).

Note: In each instance the number in parentheses corresponds to the fraction of cases in a given row. For example, in the first row 4 represents the proportion of 0.67 of the total 6 cases.

57 cases where deterrence applied, no wars resulted, while in the 242 cases in which it did not apply, 12 wars broke out. Although not overwhelming, this difference provides some support for the claim that extended deterrence works. Once again, however, the dependent variable focused exclusively on the occurrence of violent behaviour.

Geller (1990, 295–6, 300–2, 306) tested propositions focusing on the escalation of disputes and the presence or absence of nuclear weapons. He used 393 militarized disputes from 1946 to 1976 as evidence and, based on the nuclear capabilities of the initiator and defender, created the following subsets of cases: both nuclear, nuclear/non-nuclear, non-nuclear/nuclear and both non-nuclear. Geller found that nuclear capabilities did not discourage escalation by an adversary, implying that nuclear weapons might be of limited value “for achieving second-order political objectives in disputes.”

EVALUATING PROGRESS: RELIABILITY AND VALIDITY

At this point it is appropriate to identify criteria for evaluation of the above-noted studies. The most straightforward approach is to focus on validity and reliability with regard to case selection. These are standard criteria for the assessment of any aggregate analysis.

Validity

“Validity” refers to accuracy, that is, fidelity to a theory and its concepts during the process of testing. This criterion encompasses several more specific concerns related to case selection, operationalization, and research design.

First, there is a bias against reaching a conclusion that would associate nuclear weapons with successful deterrence, for the following reason. In order to engage in valid aggregate testing, it is essential to obtain a set of observed instances of threat and counterthreat. However, many situations could arise in which an attack on the client state of an adversary is deterred because the initiator is certain that the patron – a nuclear power – would retaliate effectively. The failure to threaten the client state publicly, caused by the anticipated reaction from the patron, should count as an instance of successful deterrence. Unfortunately, given the practical need to focus on observed interactions, that process would not be identified by any standard means of collecting data. While the problem of “nonevents” cannot be rectified by an approach that focuses on explicit threats, the bias against a verdict in favour of nuclear deterrence should be kept in mind when assessing aggregate findings. This particular distortion, of course, may be supplemented by the fact that a threatening state can fail to achieve its goals for reasons independent of deterrent threats, with the latter incorrectly receiving the credit. Domestic politics, to cite just one extraneous consideration, can impose significant limitations on the pursuit of foreign objectives.¹⁴

A second problem applies to attacker/defender designations, an important prerequisite for an accurate assessment of rational deterrence theory within the Huth-Russett and Lebow-Stein framework. It is often unclear, for instance, whether a decision maker initiated the threat of attack or was simply responding to what s/he perceived at the time to be a clear act of aggression. By implication, final coding decisions about success versus failure will be difficult to make. The assumption that the attacker and defender remain the same throughout a crisis is also problematic. Military-security crises often involve a series of threats and counterthreats, with each side acquiring and playing both roles at various stages in the dispute, effectively creating a “sequence of deterrence episodes” (Harvey and James 1992, 32).¹⁵ Assigning a single role to each actor in complex cases may account for some of the contradictions across data sets.

Third, several of the studies surveyed did not distinguish between immediate and general deterrence when collecting cases. This lacuna makes it difficult to assess the role of nuclear weapons, because the success or failure of general and of immediate, crisis-oriented deterrence constitute different issues. Presumably, if it focused on immediate deterrence, aggregate research would have greater validity; variables such as threat and degree of success could be measured more accurately. Unfortunately, this is no easy task. The continuing debate over case selection between Huth and Russett (1990) and Lebow and

Table 2.3
Classes of Deterrence/Compellence Encounters

<i>Crisis-Management Threat Used</i>	<i>To Deter/Compel</i>	<i>Actions Contemplated</i>
1 Nuclear		1 Nuclear
2 Large-scale war		2 Large-scale war
3 Limited war		3 Limited war
4 Minor military		4 Minor military
5 Economic		5 Economic
6 Political		6 Political
7 Diplomatic		7 Diplomatic

Note: Possible combinations of encounters are 1,1; 1,2; 1,3; 1,4; 1,5 ... 2,1; 2,2; 2,3 ... 3,1; 3,2; 3,3; and so on.

Stein (1990) suggests that even if researchers were to agree on the criteria for distinguishing cases of immediate and general deterrence (or successes and failures), there still would be very little consensus regarding the final data set. It is equally important, for conceptual and theoretical clarity, to distinguish cases of deterrence from compellence, with an understanding that a crisis may include examples of both types of threats. Huth and Russett (1990, 475) correctly pointed out that "So long as the analyst carefully singles out those aspects of behaviour and actions which conform to the definition of deterrence, then the case can be validly used to test propositions on deterrence even when the case includes examples of both deterrent and compellent behaviour." On the other hand, carefully singling out the threat-counterthreat sequence that would allow the researcher to pinpoint those aspects of behaviour that conform to deterrence or compellence is often difficult, if not impossible, to accomplish with any degree of empirical precision.

If we accepted the recommendation that we render the study of deterrence theory more "rigorous" by paying "careful attention to specifying the empirical domain within which propositions on deterrence success and failure apply" (Huth and Russett 1990, 473),¹⁶ then both the military and nonmilitary dimensions of coercive diplomacy could be subdivided even further into approximately forty-nine different classes of deterrence (or compellence) encounters listed in Table 2.3. Presumably, propositions regarding success and failure could not readily be applied across these classes either, which would represent yet another major obstacle to cumulation, given the countless coding decisions to be made prior to empirical testing of the theory.

It appears that a valid test of propositions using the success/failure framework, then, requires specification of (1) the particular class of

deterrence or compellence, (2) whether the case constituted an immediate or general deterrence or compellence encounter (or some combination), and then (3) whether the particular type of deterrence or compellence threat was a success or failure. It also requires (4) specification of who the attacker and defender were, with an understanding that officials of each state may play both roles at different points in the same crisis. The resulting complexity makes case selection and coding extremely difficult, particularly if the researcher also is required to (5) pinpoint the precise time frame and exact sequence within which the appropriately designated threats and counterthreats were made (Harvey and James 1992, 33). Although Huth and Russett (1990) and Lebow and Stein (1990) partially specified the class of deterrence they were concerned with (that is, threats of military retaliation to deter military attacks), the potential for contamination is obvious. Regardless of their efforts to focus solely on features of the crisis that appeared, on the surface, to be elements of a military-security deterrence encounter, the behavioural dynamics associated with other classes of deterrent or compellent threats still may have influenced behaviour and outcomes.¹⁷

Reliability

Reliability, the second criterion of evaluation, refers to replication of findings. When something is measured more than once, the result should not change if the measuring instrument is reliable and the purpose is the same. A straightforward method of assessing the reliability of the research program on nuclear deterrence is to compare the case lists and other coding decisions. Assuming that measurement procedures are handled competently from one instance to the next, the selection of cases should show some consistency across the studies. In other words, similar listings are taken to mean that nuclear deterrence is being investigated in much the same way from one study to the next.

Given the diversity of the studies that have produced case listings, reliability in this context should be interpreted more in terms of correspondence or convergence as opposed to measurement quality. In other words, to what extent does any given empirical domain resemble another? The higher the average level of correspondence, the more advanced the research program is considered to be.

Consistency in case selection was assessed on the basis of a coefficient of correspondence, calculated as follows. For each pair of studies, the overlapping years were identified. (As noted previously, the comparison was further restricted to cases of potential nuclear-based deterrence.)

For example, with Russett (1963) and George and Smoke (1974), the relevant years are 1948–61. In that period Russett has seven cases and George and Smoke have ten, with four in common. Two coefficients could therefore be calculated, that is $4/7$ (57.1 percent) and $4/10$ (40.0 percent), respectively. For each comparison, the higher percentage value was used, in order to present the most favourable judgment.

Table 2.4 displays the paired comparisons for the seven studies that include case lists, with the first number in each instance representing the degree of overlap (the second number is discussed in the next paragraph). These coefficients range from 27.3 to 87.5 percent, with an average of 57.8 percent. The two highest values in the table (84.6 percent and 87.5 percent) are the only acceptable ones, using the standards that apply when it is assumed that the same underlying concept is being measured. The generally low values in the table, however, reveal that the studies do not have a common vision of the process of deterrence. Instead, these coefficients highlight the variations in focus among aggregate testing efforts. One study examined deterrence failures exclusively, while another focused on the ostensible role played by nuclear weapons. Still others incorporated instances of both general and immediate deterrence in testing various propositions.

Other tests for correspondence could be instituted – Maoz (1988) provides several options – but one more should be sufficient for present purposes. Within the subset of cases common to each pair of data sets, the second number for each entry in Table 2.4 represents the degree of agreement regarding two further coding decisions. If success or failure and the roles of initiator and defender are the same for a given case, that agreement is counted as a favourable instance. When either decision has been made differently, the case is designated as one of disagreement. Because of the lack of explicit coding by Betts (1987), coefficients have not been calculated for his data set relative to the others. The coefficients that do appear range from 16.6 to 100.0 percent, with an average of 59.8 percent. The mean value is similar to that calculated for case selection, although the range is much greater. As mentioned earlier, one explanation for the discrepancy in coding success and failure and in coding roles is the potential instability in the roles played by the actors in a given case.

Keep in mind that the purpose of the preceding reliability test is to offer additional, more quantitative, confirmation of heterogeneity among studies of deterrence and to demonstrate that the prospects for cumulation are not promising precisely because there is no definitive case listing even among studies that cover the same period and type of deterrence encounter. The reliability test does not assume that

Table 2.4
Coefficients of Correspondence

	<i>Russett (1963)</i>	<i>George and Smoke (1974)</i>	<i>Organski and Kugler (1980)</i>	<i>Kugler (1984)</i>	<i>Huth and Russett (1984)</i>	<i>Betts (1987)</i>
George and Smoke (1974)	57.1/75.0 ^a	71.4/60.0	84.6/100.0	41.7/60.0	54.5	27.3
Organski and Kugler (1980)	57.1/50.0	75.0/50.0	42.9/50.0	70.0	71.4/62.0	
Kugler (1984)	42.8/66.7	54.5/16.6		70.0	33.3/50.0	
Huth and Russett (1984)	66.7/83.3	87.5	35.7/40.0			
Betts (1987) ^b	57.1	50.0/33.3				
Huth (1988a)	62.5/100					

Note: The overlapping cases in each pair are as follows: (1) Russett/George and Smoke – Berlin blockade 1948, South Korea 1950, Quemoy 1954–55, Hungary 1956; (2) Russett/Organski and Kugler – Berlin blockade 1948, South Korea 1950, Hungary 1956, Suez 1956; (3) Russett/Kugler – Berlin blockade 1948, South Korea 1950, Hungary 1956; (4) Russett/Huth and Russett – Iran/USSR 1946, Turkey-USSR 1947, Berlin blockade 1948, North Korea 1950, Quemoy 1954–55, Egypt 1956–57; (5) Russett/Betts – Berlin blockade 1948, Korea 1950, Quemoy 1954, Suez 1956; (6) Russett/Huth – Iran-USSR 1946, Turkey-USSR 1947, Berlin blockade 1948, Korea 1950, Quemoy 1954–55; (7) George and Smoke/Organski and Kugler – Berlin blockade 1948, Korea 1950, Hungary 1956, Berlin 1961, Cuba 1962; (8) George and Smoke/Kugler – Berlin blockade 1948, Korea 1950, Hungary 1956, Quemoy 1958, Berlin 1961, Cuba 1962; (9) George and Smoke/Huth and Russett – Berlin blockade 1948, Korea 1950, Taiwan 1954, Middle East 1957, Quemoy 1958, Berlin 1961; (10) George and Smoke/Betts – Berlin blockade 1948, Korea 1950, Quemoy 1954, Taiwan 1958, Berlin deadline 1958, Berlin *aide-mémoire* 1961, Cuba 1962; (11) George and Smoke/Huth (1988a) – Berlin blockade 1948, Korea 1950, Quemoy 1954–55; (12) Organski and Kugler/Kugler – Czech coup 1948, Berlin blockade 1948, Korea 1950, Hungary 1956, Berlin 1961, Cuba 1962, Vietnam 1964, Czech coup II 1968, Sino-Soviet clash 1969, Middle East 1973, China/Vietnam 1979; (13) Organski and Kugler/Huth and Russett – Berlin blockade 1948, Korea 1950, Berlin 1961, Vietnam 1964, Middle East 1967, Middle East 1973; (14) Organski and Kugler/Betts – Berlin blockade 1948, Korea 1950, Suez 1956, Sino-Soviet clash 1969, Berlin *aide-mémoire* 1961, Cuba 1962, Middle East 1973; (15) Organski and Kugler/Huth (1988a) – Berlin blockade 1948, Korea 1950, Vietnam 1964–65, Middle East 1967, China-Vietnam 1979; (16) Kugler/Huth and Russett – Korea 1950, Quemoy 1958, Berlin 1961, Vietnam 1964, Middle East 1973; (17) Kugler/Betts – Berlin blockade 1948, Korea 1950, Taiwan 1958, Sino-Soviet clash 1969, Berlin *aide-mémoire* 1961, Cuba 1962, Middle East 1973; (18) Kugler/Huth – Berlin blockade 1948, Korea 1950, Vietnam 1964, China-Vietnam 1979; (19) Huth and Russett/Betts – Berlin blockade 1948, Korea 1950, Quemoy 1954, Taiwan 1958, Berlin *aide-mémoire* 1961, Middle East 1973; (20) Huth and Russett/Huth – due to the large number of shared cases, those *not* held in common are listed: Turkey/Egypt/Syria/USSR 1957, Taiwan/China 1958, Berlin West/USSR 1961, Thailand/North Vietnam 1962, Israel/USSR 1973; (21) Betts/Huth (1988a) – Berlin blockade 1948, Korea 1950, Quemoy 1954.

^a Each cell entry refers to the percentage of cases common to the two data sets during overlapping years.

^b The rate of agreement for Betts (1987) with other studies regarding success and failure and the roles of initiator and defender has not been calculated.

scholars are trying to test the same thing on the same set of cases; it tries to illustrate that they are testing different things using different sets. Lebow and Stein's (1989b) rejection of 70 percent (38) of the 54 cases of extended immediate deterrence used by Huth and Russett (1984) is the most glaring example of such differences. Consensus was generated for only 9 percent (5) of the cases, while a majority were not even considered to be deterrence encounters by Lebow and Stein, let alone examples of successes or failures. In response, Huth and Russett (1990) created an updated list of 58 cases, but remained convinced that the original data set required only minor changes. An interesting feature of this particular dispute was that the case summaries compiled by Huth and Russett to support their coding decisions were as compelling as Lebow and Stein's (1989b, 1990).

A similar dispute over historical records occurred between Lebow (1981) and Orme (1987). Orme reviewed eight crises that were described by Lebow (1981) as examples of failures of deterrence, both as a strategy and as a theory. The debate centred around the question of whether the defender in each case failed to satisfy the requirements of an effective deterrent strategy (that is, whether the defender failed to communicate clearly the commitment, resolve, credibility, and capability to follow through on the retaliatory threat) or whether the defender met these requirements, but the initiator, nevertheless, challenged the commitment due to pressures that lie outside the framework of deterrence (for example, domestic pressures). According to Lebow and Stein (1989b, 55), "Deterrence failures that occur when resolve is unquestioned and capability potent represent failures of deterrence theory as well as of strategy." In contrast, if the defender failed to communicate to the initiator its intention to retaliate, the subsequent attack is perfectly consistent with theoretical expectations. In all eight cases, the evidence presented by Orme (1987, 98) indicated that "there were chinks (if not gaping holes) in the armour of the defender that can account for the failure of deterrence." Orme concluded that standard deterrence theory continues to provide the best guide to any statesman determined to prevent an unwarranted attack.

In the Six Day War of 1967 and the Yom Kippur War of 1973, for instance, Safran (1981), Orme (1987), and Craig and George (1990) found that both Israel and the United States failed to convince Egypt and Syria of their resolve to retaliate against any incursions into the occupied territories. According to Orme (1987, 119), Prime Minister Eshkol "publicly rejected the option of attacking Syria and urged a mutual withdrawal of forces" on 22 May 1967. This, argued Safran, "encouraged Nasser, as he himself later admitted, to believe that Israel under its existing leadership might not fight, especially if it did not

receive encouragement and support from the U.S." (1981, 397). When it became clear to Israeli officials that Nasser intended to blockade the Strait of Tiran, "Israel reacted with nothing more than verbal protests and waited nearly two weeks for the U.S. to arrange a diplomatic solution" (Orme 1987, 120).¹⁸

The contrast between this kind of talk and the reaction generally expected, particularly the appeal to the world and to an obviously paralysed United Nations to act on behalf of international order as well as Israel, could only convey the impression that Israel itself judged its strength to be inadequate to the task of picking up the gauntlet thrown down by Egypt (Safran 1981, 398).¹⁹ With respect to the Yom Kippur War, Israel was unprepared for the attack by Egypt and Syria as a result of U.S. pressures on Israel not to preempt and of the Israeli government's decision to avoid a costly early mobilization (Craig and George 1990, 188–92). Also, Sadat appeared to be convinced that the U.S. commitment did not extend to protecting territory captured by the Israelis in the 1967 war and may have questioned American resolve, given internal divisions caused by Watergate. In any case, far from representing failures of deterrence theory, as Lebow and Stein claim, both 1967 and 1973 can be interpreted as deterrence encounters in which the behaviour of both sides was consistent with expectations; the attack occurred because the defending state failed to satisfy the strategic requirements.

The case of Syria's invasion of Lebanon in 1976 is presented in the following section to illustrate the complexity associated with coding deterrence encounters. The case was selected for two reasons. First, this particular crisis was not among those included in either the Huth-Russett or the Lebow-Stein data sets, although it should have been coded as an immediate deterrence encounter. Second, the case provides an excellent illustration of how scholars focusing on separate aspects of the same crisis can provide evidence to support or refute rational deterrence theory. This case will set the stage for recommending an alternative approach towards testing in the final section.

CODING DETERRENCE ENCOUNTERS: THE SYRIAN INVASION OF LEBANON, 1976

Several points should be clarified before dealing with the Syrian invasion. First, a detailed account of either the religious groups and political factions that pervaded Lebanese society in 1976 or the relations that each had with foreign governments at the time is beyond the scope of this book.²⁰ A comprehensive account of decisions taken by the different actors in the dispute, although important insofar as

each ultimately had some impact on the Syrian invasion, would also entail a more extensive analysis than the one presented here. Only the more central actors, decisions, and events leading up to the invasion will be covered. Where appropriate, more detail regarding the specific events will be presented parallel to the narrative. Finally, the changing nature and scope of Syrian involvement in Lebanon throughout the civil war can be viewed in terms of five levels of intervention (Evron 1987, 34): (1) diplomatic intervention and pressure; (2) the supply of arms versus the threat to withdraw such supplies; (3) threats, explicit or tacit, of military intervention; (4) actual intervention with irregular units under Syrian command; and (5) intervention with regular Syrian forces. Only the latter two (the resort to force between 18 January and 1 June 1976) will be covered here, given the relevance of this behaviour to Israel's deterrent efforts.

Two important events took place in Lebanon following the outbreak of the civil war (13 April 1975): the fall of the Lebanese government and the subsequent decision by President Franjiya to resolve the crisis by appointing a military cabinet on 23 May.²¹ These developments consolidated Lebanese political factions into extreme leftist and rightist groupings and transformed the scope and intensity of the fighting between the Christian Maronite and Muslim-Palestinian coalitions (Rasler 1984; Dawisha 1978). The polarization that developed between these two factions posed a serious threat to Syria: "The possible political and social disintegration of Lebanon would undermine not only Syria's traditional perception of itself as the birthplace and guardian of Arab Nationalism but also its obstinate adherence to the often tried but seldom successful concept of Arab Unity, a principle upon which for over a decade the institutional legitimacy of the Ba'athist ruling elite has uncomfortably rested" (Dawisha 1978, 246).²² There were two outcomes that seemed probable to Syrian leaders at the time: either the Christian-Maronite coalition would decide on partition, or the Muslims (with Palestinian support) would mount a successful operation to defeat the Christians and establish a radical (anti-status quo) regime in Lebanon. Neither of these scenarios was acceptable to Syrian leaders – the first meant the establishment of a new Maronite Israel (Rabinovich 1984; Khalidi 1979), while the second would have created a beachhead against Syria for hostile regimes in Libya and Iraq. At the time, Beirut was becoming the "main haven" for opposition groups plotting to overthrow the leadership in Damascus (Ma'oz 1988, 123), so eliminating these forces certainly would have improved the position of Syria's President Asad.

In addition to the political and diplomatic problems associated with partition, Syria also faced the possibility of a military confrontation with Israel. A successful campaign by the anti-status quo coalition to

establish a radical regime (hostile to Israel) – with authority in both Beirut and southern Lebanon – clearly would have jeopardized Israel's security interest and led to military intervention.²³ The political repercussions would have been devastating to Syria, particularly in light of Asad's efforts to preserve his country's traditional role as regional protector. Occupation of southern Lebanon would also have provided Israel with a new front, thus increasing Syrian vulnerability in any future confrontation. Syrian leaders became convinced over time that a large-scale Muslim-PLO victory in Lebanon would have drawn Syria into an unwanted war with Israel.

To restore unity in Lebanon and preempt an Israeli invasion, Syrian officials attempted to mediate the dispute quickly through diplomatic channels. The day after President Franjija's unpopular decision to appoint a military cabinet, Asad sent his foreign and defence ministers to mediate the crisis (Deeb 1980, 124). As the fighting in Beirut continued, Syria proposed a political compromise in the form of a new nonmilitary cabinet headed by Rashid Karami. Given the growing opposition to Franjija's original decision, the president was forced to accept Karami, a traditional rival, as the new prime minister. On 30 June 1975 Karami formed a government of communal balance in which he also held the defence portfolio. Then, after a relatively calm period during the months of July and August, the fighting in Lebanon "escalated precipitously" as it expanded beyond Beirut – immediately after the signing of the Sinai II agreement between Israel and Egypt – and developed into an all-out civil war (Rasler 1984, 435). The Sinai agreement increased Syrian anxieties that a confident Israel might begin pursuing a more aggressive foreign policy in the North. Despite continued efforts by Syria throughout the remainder of 1975 to arrive at a political solution, the scope and intensity of the fighting in Lebanon increased.²⁴

In late winter and early spring of 1976, several events set the stage for indirect (and later direct) Syrian military involvement, the most important of which was a major defection of Muslims from the Lebanese army. The newly formed army quickly gained control of military posts in Lebanon, including several on the border with Israel (Freedman 1978, 67–9). As the Lebanese army began to weaken, the commander of the Beirut garrison declared a coup d'état, demanded the resignation of Franjija, and began moving the forces under his control towards the presidential mansion. This succession of Muslim victories posed a major dilemma for Syria; partition now appeared to be a real possibility, especially considering the pattern of fighting conducted by the Christians, who were attempting to carve a "Maronite Ghetto state" out of Lebanon (Meo 1977, 29).

In response to these events, Syrian leaders decided that some form of military intervention was required. The scope and direction of

Syrian involvement escalated dramatically throughout the winter and spring of 1976 (Dawisha 1978, 1980; Freedman 1978; Lawson 1984; Evron 1987):

- 1 Initial intervention in September 1975 was indirect and followed the signing of the Sinai II agreement. The scope of military involvement was limited to despatching a few units of the Palestine Liberation Army (PLA), under Damascus control, to Tripoli to control the fighting (Lawson 1984, 451).
- 2 In the second week of January 1976, additional PLA units (Yarmouk Brigade) were deployed in response to appeals by Lebanese Leftist leaders (Dawisha 1978, 251).
- 3 With the growing concern over the impending defeat of the Christian-Maronite forces in March and the increased likelihood of partition, Syrian decision makers ordered the PLA and units of the Syrian-backed al-Saiqa guerrilla groups to halt the Lebanese Arab Army, which began advancing towards the presidential palace.
- 4 When these efforts failed, Syria began, on March 28, to cut off supplies to the Palestinians and threatened to escalate Syrian military involvement if Muslim-PLO leaders refused to accept the reform proposals (Freedman 1984; Lawson 1984).
- 5 On 4 April, Syria imposed a blockade on Lebanese ports to prevent the warring factions from being resupplied (Lawson 1984).
- 6 One week later, regular infantry and armoured units of the Syrian armed forces moved into Lebanese territory, remaining idle until 1 June (Lawson 1984, 451).
- 7 On 1 June, the Syrian army became directly involved in the fighting, moved into Lebanon, and proceeded to engage Muslim-PLO forces in a drive towards Beirut and Sidon.

The question about this chronology that is most relevant to the present study is whether Israel's reaction to this activity constituted a valid case of immediate deterrence. Presumably, if an immediate deterrent threat was not initiated by Israel in response to Syria's invasion plans, the case would not be suitable for evaluating the theory. Once we establish that the case is appropriate, the second objective will be to illustrate how Israel's deterrent threat both succeeded and failed.

As Lebow and Stein (1989a, 53-5) correctly point out, "deterrence theory stipulates that the deterrer must carefully define the unacceptable behaviour, communicate to all possible adversaries a commitment to punish transgressors (or deny them their objectives), possess the means to do so, and demonstrate the resolve to carry through its threat (i.e., do its best to make the threat of denial or punishment credible)."

Israel's response to the initial intervention by PLA forces in September was limited to a warning by Yigal Allon, Israel's foreign minister: "As long as the quarrels in Lebanon are confined to the Lebanese community, I don't think that Israel can or should do anything. But her defense interests along the northern border may be touched if [Syria] intervenes. In such a case, Israel should preserve her moral right and her military might to protect her security interests in the border zone" (quoted in Evron 1987, 36-7). A second warning, several weeks later, was delivered by Prime Minister Rabin: "Israel has an interest in the existence of Lebanon in its present political form. Any attempt to take over Lebanon will constitute a situation which adversely affects our security ... Israel cannot remain indifferent in the event of foreign intervention" (37). Similar statements were made by Israeli leaders throughout the winter and early spring of 1976.

In addition to the distinction drawn between regular units and the PLA, Israeli officials communicated to Syria the critical importance of its security interests in southern Lebanon and also began to distinguish acceptable weapons systems: tanks and surface-to-air missiles (SAMS), for example, were considered unacceptable. Finally, along with verbal statements and public announcements, Israel adopted two additional channels of communication for its deterrence signals: messages were delivered to Syria via Washington that stressed Israeli concern about direct Syrian deployment; and military moves were initiated by Israel around the northern border with Lebanon.

In March, Israel communicated to Syria the final set of conditions, which were to become known as the "red lines."

- 1 Syria should not intervene in a declaratory (open) way.
- 2 The intervention force should not exceed one brigade.
- 3 The force should not employ tanks, artillery, or SAMS.
- 4 Syrian aircraft should not operate in Lebanon, nor Syrian naval units in Lebanese waters.
- 5 Syrian units should not move south of the line running ten kilometres km south of the Beirut-Damascus highway.

It is clear from these warnings that Israel satisfied each of the conditions stipulated by standard deterrence theory. The leaders carefully defined the unacceptable behaviour, communicated their commitment to Asad, and possessed the means to punish Syria and deny Syria its objectives if the red lines were violated. Although Israel did not specify in detail the nature of the retaliation, Israel's previous military activity in the region was sufficient to establish its resolve to carry through with its threat.

Now, if the conditions for an immediate deterrence encounter were satisfied, but the proscribed behaviour still occurred, that would constitute a case of deterrence failure – both in terms of theory and strategy (Lebow and Stein 1989a). In this case, the evidence indicates that Israel's deterrent threat was challenged, given Syrian violations. Notwithstanding the red lines, Syria's intervention consisted of 6,000 men (almost three times the size of one brigade, the limit stipulated by Israel) and 160 tanks (160 more than the limits stipulated). In addition, the Syrian navy was used off the coast of Lebanon on 4 April to establish a blockade and prevent the warring factions from being resupplied. All things considered, although Israel's resolve was unquestioned and its capability potent, Syrian leaders clearly went beyond the red lines.²⁵ Several scholars have produced an impressive body of evidence to explain why Syrian leaders challenged Israel's deterrence efforts in 1976, but limitations on space prohibit a more detailed analysis here.²⁶ Based on the evidence outlined above, it appears that deterrence, both as a theory and as a strategy, failed.

There is, however, an alternative interpretation of the case that is consistent with theoretical expectations; perhaps Israel's commitment or resolve to follow through with its retaliatory threat was not clearly communicated to Syria. If the commitment was ambiguous, then Syria's actions would be consistent with theoretical expectations and would not constitute anomalies that jeopardize the validity of rational deterrence theory.

In addition to stipulating the red line conditions, Israel communicated to Syria that retaliation for violations of conditions 1–4 would be limited to strategic assets linked to the security of northern Israel, without any real threat of punishment beyond that (Evron 1987). If Syrian forces were to violate condition 5, however, Israel's military response would be far greater. But even this condition was further qualified – although such violations would provoke a more serious military retaliation, a direct strike against the Syrian forces was never identified as an option. It was reasonable for Asad to expect, therefore, that an invasion supporting the Christian-Maronite coalition would not provoke an Israeli counteraction, as long as the most important of the conditions was not violated.

Had Israel established clearer (more precise) guidelines and communicated to Syria the military repercussions of each violation, the deterrent threat would likely have been more effective, but Israel would have been forced to respond (in theory) to any move by Syria outside the strict confines of the red lines, a very risky threat. Also, Syria would have lost the window of opportunity to respond to the Lebanese crisis in a way that was perfectly consistent with Israel's

security objectives at the time. In other words, the outcome of a successful deterrent threat was less appealing to Israeli strategists, primarily because Israel stood to gain from a limited Syrian invasion into Lebanon: they could ensure military support for the Maronites, who were on the verge of defeat, without getting involved in a confrontation; avoid jeopardizing their relations with the United States; be certain that a Syrian intervention would affect Arab unity; and promote a battle between Syrian forces and the PLO.

In order to render sound judgments about deterrence success or failure in the Lebanese crisis, therefore, a distinction must be drawn between Israel's intrinsic and strategic interests (Jervis 1979). Although Syria was deterred from intervening in a way that was contrary to the intrinsic interests of Israel (its survival, territorial integrity, and so forth), Syria still invaded in a way that challenged many of Israel's "strategic" interests as stipulated in the red lines. Unlike intrinsic interests, which are rarely questioned, "judgments about the credibility of strategic interests may turn on efforts a defender makes to convey its resolve" (Lebow and Stein 1989b, 63). If success is measured, at least in part, on the basis of whether the strategic objectives of the deterrer are satisfied, then Syria's invasion should be considered a partial deterrence failure; Israel's efforts to convey its commitments and resolve were restricted (intentionally or not) and provided Syria with a window of opportunity to move into Lebanon. On the other hand, if the outcome is assessed in terms of Israel's intrinsic interests, the case should be classified as a deterrence success, both in terms of theory and strategy, because the most important of the five conditions (the geographical line ten kilometres south of the Beirut-Damascus highway) was not violated. Depending on which of the red line conditions one highlights and the specific time frame one selects, the final coding of this case will change.²⁷

SUMMARY

As the case just discussed demonstrates, identifying cases appropriate for evaluating deterrence theory within the Huth-Russett/Lebow-Stein research program is problematic. Although the authors' efforts have been constructive, there are numerous problems with the success/failure framework that cannot easily be overcome through reference to the historical record. Based on the preceding analysis, new directions for aggregate testing of nuclear deterrence are presented in chapter 3. The alternative approach uses international crisis data to test several new propositions derived from deterrence theory, rational choice theory, and coercive diplomacy more generally. The most

important difference between the two testing strategies is that coding of failed and successful deterrence, identifying attackers and defenders, differentiating deterrence from compellence, and distinguishing immediate, general, and extended dimensions of the strategy all become peripheral issues, thus avoiding many of the pitfalls and associated debates noted earlier. Instead, the emphasis shifts to provocation and retaliation by leaders facing direct threats to fundamental values, a finite time for response, and a high probability of involvement in military hostilities.

Given the focus on rivalry and bargaining during intense crises, the approach entails a less demanding set of criteria for selecting cases appropriate for testing both conventional and nuclear deterrence theory. This is especially important with respect to the latter issue, because the literature on nuclear deterrence is primarily speculative and focuses on the behaviour expected from nuclear antagonists faced with the prospect of mutual annihilation. By comparison, insufficient attention has been paid to the relationship between expected behaviour and actual results.

3 New Directions for Aggregate Testing

Two basic tasks have been accomplished in chapter 2. In accordance with the recommendations detailed assessment of the research program on nuclear deterrence identified areas of disagreement and consensus. Based on criteria of reliability and validity, the findings suggest that aside from dealing with deterrence in one form or another, the program does not have a common rationale and, as a consequence, has not cumulated. The question of whether states act according to the logic derived from standard applications of deterrence theory, particularly in the nuclear realm, remains unanswered. Taking into account the accomplishments and shortcomings of the dominant testing strategies, this chapter outlines new directions for testing. It begins by describing the coding procedures and selection criteria for twenty-eight cases of superpower rivalry that serve as the empirical base for the remainder of the investigation.¹ The debate over rational deterrence – especially at the nuclear level – cannot move forward without an authoritative body of evidence.² The intention, therefore, is to provide a more clearly identified and common purpose in order to facilitate convergence of the research program.

CASE SELECTION AND CODING PROCEDURES

There are several propositions embedded within nuclear deterrence that are related to rational choice and coercive diplomacy more generally and that can be evaluated using a crisis-based data set, rather than one that focuses solely on cases of immediate deterrence. The

most important difference is that coding of failed or successful deterrence, differentiating deterrence from compellence, and distinguishing immediate, general, and extended deterrence, all become peripheral issues. The emphasis now is on provocation and retaliation by superpowers experiencing direct threats to fundamental values. Given the focus on rivalry and bargaining between nuclear powers in intense crises, and the more basic issue of the role played by nuclear weapons in that process, the approach entails a less rigid, but no less valid, set of criteria for selecting cases appropriate for testing nuclear deterrence. Focusing exclusively on cases of superpower rivalry during foreign policy crises also eliminates the need to specify the existence or absence of specific kinds of deterrent or compellent threats, thus avoiding many of the pitfalls of previous work on the subject.

Before turning to the subset of superpower crises, it is appropriate to describe case selection and coding procedures of the International Crisis Behaviour (ICB) data set in general terms. Initial identification of cases (by two coders) relied upon the *New York Times Index* and *Keesing's Contemporary Archives*. Area experts and existing compilations of "related phenomenon" were consulted in the process of compiling a final list (Wilkenfeld, Brecher, et al. 1988, 15; Brecher, Wilkenfeld, et al. 1988, 32). Intercoder reliability was 0.85.

For a given state, a foreign policy crisis arises when its central decision makers perceive three interrelated conditions: (1) a threat to basic values, with a simultaneous or subsequent awareness of (2) finite time for response and of (3) the high probability of involvement in military hostilities (Wilkenfeld, Brecher, et al. 1988, 2). Thus, in all cases, a specific danger is experienced by an actor who seeks to respond to some tentative change in the status quo.

The incorporation of perceived finite time for reaction represents a departure from previous treatments, effectively putting less arbitrary boundaries on cases of intense crisis. Table 3.1 displays a set of cases that meet the requirements derived from the review in chapter 2 and that are appropriate for pursuing the alternative research agenda described below. Each entry in the table meets the definition of a foreign policy crisis developed by the ICB project. It should be noted that foreign policy crises, which are experienced by individual actors, take place within the larger setting of an international crisis as described by Brecher, Wilkenfeld, et al. (1988, 3). An international crisis is a situational change that produces disruptive interactions and challenges the existing structure of the international system. In other words, it is identified objectively. Within the boundaries of an international crisis are the individual experiences of the actors, some of which are relevant to the current analysis. The international crisis over the Six Day War, for example, involved actors other than the superpowers,

Table 3.1
 Superpower Rivalry in Foreign Policy Crises, 1948–88

<i>Case</i>	<i>Provocation by^a</i>	<i>Retaliation by</i>	<i>Trigger Date^b</i>	<i>Termination Date</i>
1 Berlin blockade	USSR	USA	24/06/48	12/05/49
2 Berlin blockade ^c	USA	USSR	07/06/48	12/05/49
3 Korean War I	USA	PRC/USSR	26/06/50	-/07/50 ^d
4 Korean War II	USSR/PRC	USA	31/10/50	10/07/51
5 Korean War II ^e	USA	USSR	07/10/50	03/01/51
6 Korean War III	USA	PRC/USSR	22/05/53	27/07/53
7 Taiwan Straits I	USA/Taiwan	PRC/USSR	-/08/54	-/11/54
8 Taiwan Straits I	USA/Taiwan	PRC/USSR	02/12/54	23/04/55
9 Suez-Sinai Campaign	USSR	USA/Br/Fr	05/11/56	08/11/56
10 Syria-Turkey	USA/Turkey	USSR/Syria	07/09/57	29/10/57
11 Taiwan Straits II	USA/Taiwan	PRC/USSR	27/08/58	30/09/58
12 Berlin deadline	USSR	USA/Br/Fr	27/11/58	30/03/59
13 Berlin deadline	USA/Britain/France	USSR	15/12/57	15/09/59
14 Bay of Pigs	USA	USSR/Cuba	15/04/61	24/04/61
15 Berlin Wall	USSR/GDR	USA/FRG	13/08/61	17/10/61
16 Cuban missiles	USSR/Cuba	USA	16/10/62	20/11/61
17 Cuban missiles	USA	USSR/Cuba	22/10/62	20/11/61
18 Congo I	USA/Belgium	USSR	24/11/64	17/12/64
19 Gulf of Tonkin	USA	USSR/N.Vietnam	04/08/64	07/08/64
20 Six Day War	USSR	USA/Israel	06/06/67	11/06/67
21 War of Attrition	USSR/Egypt	USA/Israel	19/03/70	07/08/70
22 Cienfuegos base	USSR	USA	16/09/70	23/10/70
23 Yom Kippur War	USSR/Egypt/Syria	USA/Israel	12/10/73	31/05/74
24 Ogaden III	USA	USSR/Ethiopia	30/06/82	-/08/82
25 Invasion of Grenada I	USA	USSR/Grenada	22/10/83	28/10/83
26 Invasion of Grenada II	USSR/Grenada	USA	19/10/83	28/10/83
27 Nicaragua, Mig 21s	USA	USSR/Nicaragua	07/11/84	12/11/84
28 Nicaragua, Mig 21s	USSR/Nicaragua	USA	06/11/84	09/11/84

^a "Provocation by" refers to the first superpower to become involved in the crisis. Any action taken by a coalition that included a superpower was coded accordingly (e.g., the United States, Britain, and France in the Berlin blockade).

^b The trigger and termination dates refer to the range of involvement for the responding state as a crisis actor. For example, the United States perceived a crisis on 13 August 1961 when confronted with the creation of a wall between East and West Berlin. The United States ceased to meet the conditions of a crisis actor on 17 October 1961 when a speech by Khrushchev entrenched the status quo in Berlin.

^c Some international crises contain more than one relevant foreign policy crisis. For example, the Berlin blockade encompasses two actor-level cases, with first the USSR and then the United States in the position of responding to a perceived threat from its adversary.

^d The cases appear in order of trigger date. When a dash appears instead of a number, this means that a specific date of trigger or termination could not be specified on the basis of available sources.

^e Some of the cases take place within the boundaries of an ongoing war. As opposed to the war as a whole, the relevant aspect is the superpower interaction. Thus, Korean War II focuses on the United States' reaction to the threat posed by China's expanding role in the conflict in conjunction with Soviet backing.

but the pertinent aspect of that case is the reaction by the United States to a perceived challenge from the USSR. There may also be more than one foreign policy crisis that is relevant to the present study within a given international crisis. More complex cases, such as the Berlin blockade, Taiwan Straits I and the Cuban missile crisis, appear in component form, reflecting changes in roles and new threats during the full span of an international crisis. In all the foreign policy crises listed in Table 3.1, a provocation has been perceived by a superpower and attributed to its rival.

There are at least three reasons why superpower rivalry is emphasized instead of immediate deterrence. First, the threat-based criterion for case selection helps to bypass the need to distinguish different forms of coercive interaction (that is, compellence or deterrence). The Berlin deadline crisis, for example, might be construed as an attempt by the United States to compel the USSR to withdraw its announcement of a deadline for signing a peace treaty with the German Democratic Republic (GDR). The latter action, which would have granted further recognition to the GDR and undermined the Western powers' position in Berlin, could also be placed in the context of deterrence: the United States and its allies hoped to deter the proposed signing of a Soviet-GDR peace treaty. To some extent, as in this example, designation of a given case as deterrence or compellence is a matter of emphasis (Huth and Russett 1990). The threat-based criterion for selection reduces some of the ambiguity associated with deciding that issue.³

Second, although not essential, it is appropriate to examine cases that involve the USSR and the United States as adversaries. In other words, it is appropriate to examine cases in which one superpower is threatened directly or indirectly by the other, as opposed to situations of general deterrence like those explored by George and Smoke (1974) or Kugler (1984). This criterion eliminates intrabloc crises, such as the Hungarian crisis in 1956, in which one superpower threatens its client state, not the more immediate interests of the rival. As for unilateral nuclear deterrence or compellence, such cases are regarded as a separate class of events (Geller 1990). Furthermore, to ensure that nuclear weapons have at least a potential role to play, it is appropriate to exclude from consideration those conflicts in which the client state of one superpower threatens the client state of the other. Many of the dyads that Weede (1981, 1983) used, for example, would be ruled out by this criterion. In such conflicts it can be virtually impossible to attribute responsibility to a superpower.

Israel, for example, experienced a crisis in its relations with Egypt on 10 April 1973, when intelligence reports suggested that an Egyptian

attack would occur on 15 May (Brecher, Wilkenfeld, et al. 1988, 333–4). At that time Egypt and Israel could be regarded as clients of the USSR and the United States, respectively. Yet the Israeli leadership could not have known whether the potential assault had been prompted by Moscow, despite the latter's close connection to Cairo. Furthermore, had the anticipated attack occurred, Israel and the United States still could not have been certain that the attack had – or had not – been encouraged or dictated by Moscow. Such crises, in short, make it difficult to test models of nuclear deterrence because the responsibility for initiating hostilities cannot be attributed clearly to one of the superpowers. Given the lack of authoritative evidence about decision making even in some very prominent cases (Stein 1985, 38), exclusion is prudent.

Third, in the cases in Table 3.1 a superpower occupied the role of crisis actor and the rival involved triggered the crisis. With the superpowers as adversaries it is feasible to assess the practical involvement of nuclear forces in determining the outcome. Data are available on the crisis-management techniques adopted by participants in each case, thus permitting an evaluation of the role of nuclear weapons in the process of bargaining. Consequently, the proposed case listing begins with crises underway in 1949.⁴ During that year the USSR demonstrated its nuclear capacity somewhat earlier than expected by authorities in the West. Thus 1949 is considered to be the first year in which both the United States and the USSR could have issued nuclear-based threats. Of course, only a remote chance existed that the USSR could have delivered even a single weapon to American soil, possibly through a one-way bomber mission. The USSR did, however, have some potential to strike at Western Europe; at that time, such a strike would almost have meant an assault upon the United States itself. Also, although the United States had an advantage in numbers, it had somewhat limited delivery potential. As Dingman (1988, 52–3) noted, “Despite flaws in enemy delivery capabilities, the grim truth was that Moscow’s ability to strike the American heartland was growing ... Washington acknowledged real limitations in America’s ability to put nuclear weapons on enemy targets ... Strategic Air Command planners estimated that it would take three months to bomb Moscow into submission, given the inadequacy of forward bases and overseas fuel supplies.” The USSR was becoming more of a threat, not only because it had successfully tested the bomb but because American officials began to acknowledge some of their own limitations. Thus, the Berlin blockade crisis is an appropriate point of departure for assessing the behavioural dynamics of nuclear rivals.

SUPERPOWER RIVALRY AND NUCLEAR
DETERRENCE

ICB case listings of superpower crises cannot provide a definitive account of perceptions and intentions and can offer only a relatively accurate account of actions and reactions. Consequently, critics might point to the lack of information on subjective costs, benefits, and probabilities as an important problem, especially if the goal is to test propositions derived from deterrence theory. But leaders are constrained and encouraged by military, political, domestic, and systemic forces when making foreign policy and base their subjective estimates of costs and benefits on these concrete considerations; they do not simply pull utility estimates out of the air.⁵ To imply that these factors are not important assumes that decisions are a product of internal psychological beliefs, perceptions, and intuition alone, with little, if any, connection to objective reality. If this assumption is true, then we should not expect one theory to predict behaviour any better than another, given the idiosyncratic nature of military and political calculations during crises. On the other hand, if behaviour appears to be consistent with theory, that result would, in itself, be an important finding. That is to say, for purposes of testing nuclear deterrence theory, I assume that if the probability of a retaliatory threat was high and the proscribed behaviour (for example, escalation of violence or hostility in a crisis) did not take place, then that result would constitute strong evidence that the threat had something to do with behaviour. The onus would be on critics to provide counterevidence that the threats were insignificant or offer a more compelling explanation for observed behaviour. Some scholars have argued that cases of superpower rivalry in which a crisis erupted actually constitute anomalies for nuclear deterrence theory. But this argument assumes that there were no grey areas during the Cold War, that each side understood and accepted the other's sphere of influence, and that global commitments were clear and unambiguous. If these assumptions are inaccurate (and they are) we would expect challenges to occur and crises to erupt. But these challenges do not necessarily constitute failures of deterrence or anomalies from the perspective of the theory, because it is the escalation process within these crises that is the key to determining the utility of deterrence, both as a theory and as a strategy.

Other sceptics might demand the use of primary documentation before making valid inferences about intentions and deterrence outcomes. After all, as noted in chapter 2, defenders (or challengers) may never have intended on retaliating (or attacking) in the first place,

and without access to the decision-making record it would be difficult to make those judgments about intentions. There are a few problems with this claim that should be noted. First, decision makers may not always understand their own motives or be able to articulate them (Huth and Russett 1990). Alternatively, they may be prone to articulating conflicting intentions, disguising them, or even changing their motives during official meetings. It would be a mistake, therefore, to assume that the decision-making record is essential for making valid inferences about deterrence theory.

Also, several studies have used information and data, in one form or another, obtained from primary documentation (for example, Huth and Russett 1984, 1988, 1990; Lebow and Stein 1989a,b, 1990; Harvey and James 1992, 1997) yet all have produced divergent case histories; have disagreed over the decision makers' knowledge, motives, perceptions, and intentions; and have drawn different conclusions about rational deterrence theory and strategy. Although the decision-making record is important, the record of actions, threats, counterthreats, and other forms of *observable* behaviour (as reported in the ICB listings) is equally important when evaluating predictions derived from deterrence, or any theory of coercive diplomacy, for that matter. It certainly is true that the definitive link between theory and evidence will remain elusive without access to the minds of decision makers, but it would be a mistake to conclude that we are somehow prohibited from making inferences based on observable behaviour simply because, on occasion, actions may not accurately represent intentions. This cannot possibly be true in every case, and the margin of error is likely to be sufficiently small to justify an approach that downplays that possibility. The alternative is to fully accept the inadequacy of such an approach, throw one's hands up in failure and repudiate the entire enterprise. The relevant question is whether we would know more or less about crisis management and nuclear deterrence if we pursued that strategy.

Finally, identifying cases of superpower crisis does not require the same attention to detail that is necessary when classifying immediate deterrence encounters within the Huth-Russett/Lebow-Stein framework. Any immediate deterrence confrontation constitutes an international crisis by almost any definition, but it is much more difficult to establish whether a specific crisis is an immediate deterrence encounter, given the multiple avenues available for scholars to accept or reject the coding of any one case. The "crisis" approach recommends using a less rigid, but no less valid, set of selection criteria that is equally relevant for testing a set of alternative propositions derived from theory. Critics might question whether, in the end, we can avoid the

various issues having to do with deterrence encounters that divide Huth and Russett and Lebow and Stein, but the more interesting (and relevant) question is how we can test propositions of deterrence theory if we cannot get beyond the thorny issues and coding controversies that divide Huth and Russett and Lebow and Stein. Although debates over the accuracy of historical accounts are constructive, lingering divisions over coding of deterrence successes and failures have become counterproductive. This is precisely why more effort needs to be directed towards developing alternative strategies that lie outside the success/failure framework and testing a wider range of propositions.

Analysis of superpower rivalry, then, reflects a more basic concern for the interaction between empirical research and the building of theory related to nuclear deterrence. While there has been extensive empirical work on conflict and cooperation in Soviet-American relations (Patchen 1990), few have looked specifically at game theory and its application to crisis situations to obtain a reliable judgment of nuclear deterrence. The first of two testing strategies to be described below is based on this approach.

More specifically, chapter 4 examines propositions about preferences and payoffs embedded within four alternative game-theoretic models of strategic choice: tit-for-tat (TFT) (Axelrod 1984); optimal threats (OT) (Brams and Kilgour 1988); graduated and reciprocated initiatives in tension reduction (GRIT) (Osgood 1962; Lindskold, Betz, and Walters 1986); and the converse of GRIT, referred to as GRIM (Mueller 1987). The focus will be on patterns of response when one superpower is faced with a crisis resulting from a threat attributed to the other (James and Harvey 1992). The question these models address is whether a retaliatory threat during a crisis involving two nuclear opponents will be more successful if it is proportional to the level of threat (TFT), more than proportional (GRIM), less than proportional (GRIT), or a mixture of all three (OT). Each suggests an alternative strategy for achieving the optimal outcome in a crisis and reveals the diversity of game-theoretic models that is generally overlooked by critics of rational choice. Once again, the neglected question in the literature is this: Which of the models, if any, provides the most accurate depiction of the behaviour of superpowers during intense crises? A second and equally important objective is to determine whether approximation to one strategy or another can explain crisis outcomes (for example, relative success in achieving objectives). The goal is to identify and test propositions about preferences and payoffs embedded within each model. The evidence will help determine whether the United States and the USSR follow the logic associated with classical nuclear deterrence theory or whether their behavioural patterns are inconsistent with theoretical expectations.

The second approach, developed in chapter 5, extends the analysis of nuclear deterrence by testing two additional propositions central to the theory. The first proposition deals with the relationship between the severity of the threat that, according to the theory, is expected to restrain conflict behaviour and the level of violence experienced by superpowers in crises. Although studies by Bueno de Mesquita and Riker (1982), Kugler (1984), and Betts (1987) have been informative, empirical evidence concerning the link between nuclear threats and crisis escalation remains insufficient. Of course, despite the importance of this proposition, deterrence means much more than simply avoiding war or controlling escalation. Rather, "deterrence is successful when nations that hold nuclear arsenals preserve essential policy objectives by the threat of nuclear devastation" (Kugler 1984, 474). The proposition that nuclear states preserve policy goals in disputes with non-nuclear states and achieve draws when confronted with other nuclear powers also requires testing in order to see whether the observed behaviour of the superpowers is consistent with the assumption of rationality implicit in this argument, that is, whether interactions among states confirm the logic derived from nuclear deterrence theory.⁶

The data we require to accomplish this testing are available on several outcome variables for each case listed in the Table 3.1. In order to test these propositions, therefore, it is important to focus on cases for which reliable judgments can be made regarding the attainment of objectives by nuclear-endowed participants. The ICB data set includes variables that focus on specific goal achievement and general satisfaction with the outcome as perceived by crisis actors (Brecher, Wilkenfeld, et al. 1988). The severity and centrality of violence – more traditional indicators of the degree of success achieved by the defender – are also available in the ICB holdings. Thus the data set allows for testing of the general utility of nuclear weapons, by linking crisis management techniques to goal achievement, and of the more traditional notion that possession of nuclear weapons will be connected to the limitation of violent behaviour by a threatening power.

4 Game Theory and Superpower Rivalry: 1948–88

Most applications of game theory are descriptive in nature and focus primarily on the rules of a game, choices available to the players, payoff structures, whether the issue in question is more suitable for iterated or noniterated games, and so on. The usual objective is to identify the appropriate analogy for a specific situation and derive strategies to achieve the best outcome, given the payoffs. Axelrod (1984), Lindsfold, Betz, and Walters (1986), Mueller (1987), and Brams and Kilgour (1988) are four examples, each providing an alternative strategy for achieving the best outcome in a mixed-motive interaction.

In a similar vein, a portion of the game-theoretic work is directed toward “solution concepts” – that is, theories of choice – that focus on predicting payoff outcomes in N -person, side-payment games (Michener, Yuen, and Geisheker 1980; Michener and Potter 1981; Shubik 1986). These games are analytical extensions of the traditional “chicken” and “prisoner’s dilemma” (PD) games, based largely on theories of rational choice as they pertain to cooperation, bargaining, and distributional strategies in the economic realm. The objective is to develop and test models that predict payoffs from cooperative agreements and the distribution of gains among the participants. Since this particular dimension of game theory does not directly address the conflict/cooperation transition in terms of security issues, it will not be discussed further.

Others approach the subject from the “top down,” treating game theory as a deductive model of international relations (Oye 1985;

Axelrod and Keohane 1985; Snidal 1985; Beer 1986). According to one proponent, Snidal (1985, 27), this approach requires giving empirical content to game theory "through its assumptions (e.g., about preferences and payoffs) rather than just adapting the theory (or one of its many models) to fit some historical or current event."

This chapter aims, in part, at connecting the first and third approaches. Rather than simply adapting game theory to fit superpower rivalry (post-World War II), the goal here is to identify and empirically test assumptions about preferences and payoffs derived from four game-theoretic models of strategic choice. The objective is to assess these models as descriptions of the behaviour of superpowers faced with foreign policy crises. Introducing an empirical dimension to formal modelling is expected to facilitate understanding of nuclear deterrence in particular and international relations in general.

With these priorities in mind, this chapter unfolds in six stages. First, each of the models is presented in abstract terms. A diagrammatic exposition appears in the second stage. The third task is to operationalize the strategies, with an emphasis on identifying options that might be pursued by a superpower in a crisis resulting from a threat attributed to its rival. Fourth, data from the ICB project (as described in chapter 3) is used to test the predictive power and general relevance of the models. The significance of the results is summarized in the fifth stage. The sixth and final task is to assess the overall cumulative-ness of the research by comparing the approach to similar efforts by Goldstein and Freeman (1990) based on "events data."¹

FOUR GAME-THEORETIC MODELS OF STRATEGIC CHOICE

Among the strategic options the most prominent is tit-for-tat (TFT). It was proposed by Rapoport and Chammah (1965) and elaborated by Axelrod (1984). Several interrelated questions provided the focus for Axelrod's extensive treatment of TFT: "First, how can a potentially cooperative strategy get an initial foothold in an environment which is predominantly noncooperative? Second, what type of strategy can thrive in a variegated environment composed of other individuals using a wide diversity of more or less sophisticated strategies? Third, under what conditions can such a strategy, once fully established among a group of people, resist invasion by a less cooperative strategy?" (1984, viii).

Through repeated, bilateral play in a computerized tournament, TFT achieved the highest aggregate score among the many strategies entered. In its simplest form, TFT refers to cooperation and proportional

response. The strategy proposes cooperation on the first move and a matching response on subsequent rounds of play, that is, cooperation or noncooperation, depending on the opponent's move. Axelrod (1984, 109–23) pointed to four essential properties of a stable, cooperative strategy in a durable, iterated PD game: niceness, provocability, reciprocity, and forgiveness. Niceness required that interactions begin with a cooperative move. Success was based on retaliating only when provoked by the opponent's defection, and reciprocating defection as well as cooperation (that is, success was based on provocability). Finally, forgiveness implied that a successful strategy should not continually punish a past defection but it should return to cooperation once an appropriate retaliatory move had been played.

Axelrod's claims about the success of TFT have been challenged on several grounds. For instance, since there are various player goals – not only maximization of payoffs – different strategic approaches are required to accommodate them. In more precise terms, Behr (1981, 289–300) noted that “individual and group rationality are neither entirely mutually exclusive, nor wholly compatible. There are occasions when defeating the opponent is more important than maximizing one's own payoff.” In these situations, TFT is not the optimal strategy; it is best only when the primary concern is for the welfare of both adversaries. Behr used the same tournament-based data in a different conceptual environment, ranking wins and losses (victory and defeat), not merely the achievement of goals by one player. Of course, Behr's criterion may be somewhat problematic in the context of post-World War II nuclear rivalry. What does outright victory or defeat mean in such a competition? It may be an unattainable goal.

As an alternative to TFT, Brams and Kilgour (1987a,b, 1988) developed a model of crisis bargaining based on the game of chicken.² Among their central goals was to isolate the optimal threat (OT) to cope with observed or potential provocation in a two-person conflict game. Ultimately, such a pattern of response would be expected to transform a relationship of conflict into one of cooperation. The basic question posed by Brams and Kilgour was “under what conditions should the threatened level of retaliation be less than proportionate, proportionate (i.e., tit-for-tat), or more than proportionate to deter aggression?” (1988, 104). The answer depends on the payoffs associated with the various outcomes. If the purpose, argued Brams and Kilgour, is to deter aggression at minimum costs “one must tailor the threatened punishment to the level of provocation or aggression.”

To illustrate the point Brams and Kilgour used a more sophisticated version of chicken, referred to as the “threat game.” The components of the threat game are as follows (1988, 105):

- 1 The players can make quantitative choices of levels of cooperation (c) or noncooperation (c'), not just qualitative choices of c or c' .
- 2 Once these initial choices, which can be interpreted as levels of nonpreemption (versus preemption), are made, only the less preemptive player (that is, the player who chose the lower level of preemption initially, if there was one) can retaliate by choosing a different, and presumably higher, level of noncooperation subsequently.

There are two ways in which the game can terminate. If the initial levels of c or c' are the same, play is terminated; when they are not equivalent, the game ends after the more cooperative (that is, less coercive or preemptive) player has retaliated. Thus Brams and Kilgour (1988, 105) have telescoped escalation into a "single retaliatory countermove" by the less preemptive player. Further rounds of escalation could be included, but the present simple sequence conveys the "process that might trigger further escalation" and the basic meaning of deterrence: "averting conflict through the threat of retaliation, which, if carried out, could be costly to both players." Played on a unit square, the game appears as Figure 4.1. The pure payoffs in the corners correspond to those in the chicken matrix, with r_i = payoff to Row and c_j = payoff to Column ($i, j = 1 \dots 4$). The subscripts provide the payoff rankings, with 4 as the best and 1 as the worst. The points along the x - and y -axes combine to cover all possible payoff combinations to the players.

The players, Row and Column, select preemption levels s and t , respectively. These choices range from 0 (maximum preemption) to 1 (minimum preemption). (The pure outcomes in the respective corners of Figure 4.1 are based on the combinations of polar strategic points. For example, (0,0) represents mutual escalation to nuclear war, with corresponding payoffs (r_1, c_1) .) Strategic response by the players determines a point on the unit square, referred to as the *initial position*. The *preplay position* is represented by the strategy pair (1,1), prior to noncooperation by the players. It produces (r_3, c_3) , the second-best payoff for each player. Movement from the preplay position to the initial position means that the players go from (1,1) to noncooperation at certain levels, resulting in (t,s) , or $(1/2, 2/3)$ in Figure 4.1. In this example, the more cooperative player, Row, retaliates with $q(t) = 1/3$. (If Column had been the initially more cooperative player, its response would have appeared as $p(s)$.) The *final position* in the example therefore becomes $(1/2, 1/3)$.

Brams and Kilgour then asked the following basic question: "For each possible level of aggression or preemption, what minimum level

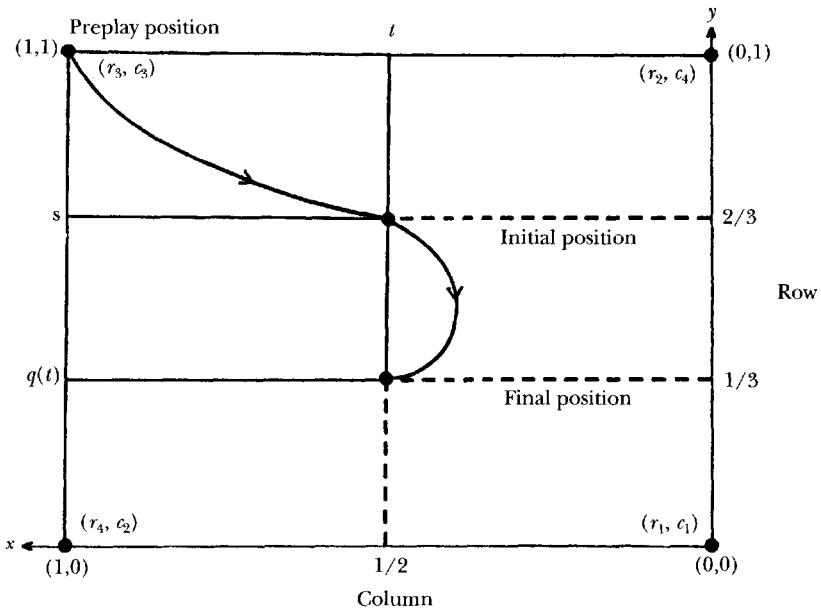


Fig. 4.1 The threat game

Source: Brams and Kilgour (1988, 105).

Key: (r_i, c_j) = (payoff to Row, payoff to Column).

r_1, c_4 = best; r_3, c_3 = next best; r_2, c_2 = next worst; r_4, c_1 = worst.

s, t = initial strategy choices of Row and Column, respectively.

$q(t)$ = subsequent strategy choice of Row (more cooperative player initially).

of retaliation will be required to deter a player from preempting?" The precise quantitative levels and crossover points are based on the retaliation-to-provocation (RP) ratio. More specifically, the RP ratio measures "the relative deviation from proportionality, or tit-for-tat, that just deters a potential aggressor." In terms of the crossover points, with increasing intensity of an act represented by a larger number "an RP ratio greater than 1 would indicate that one's retribution must be more punishing than an opponent's aggression, whereas a ratio of less than 1 would indicate that one's retribution can be less punishing than the aggression" (1988, 113). They concluded that, in at least some circumstances, a *more* than proportionate response is optimal when facing relatively minor aggression, and a *less* than proportionate response is optimal for deterring relatively major aggression.

Although both Axelrod and Brams and Kilgour used game theory to identify an ideal strategy, they differed with respect to which of the two standard analogies – chicken or PD – is more realistic in portraying interstate conflict. Theorists generally agree that chicken is a more

appropriate analogy for *strategic nuclear* rivalry, given the penalty of mutual noncooperation – nuclear devastation. PD, on the other hand, is considered more analogous to conventional deterrence or an arms race.³ Since each model focuses on a different type of strategic choice, it may be more appropriate to view them as complementary.⁴ Notwithstanding the discrepancies between Axelrod and Brams and Kilgour, however, there is an overriding similarity: each model represents an effort to identify the best strategy for transforming conflict into cooperation. Brams and Kilgour dealt with this transition in terms of nuclear-based bargaining, while Axelrod focused on strategic interaction in a more general sense. In terms of this transition, therefore, it would not be inappropriate to view the strategies as alternatives.

Osgood (1962) offered a programmatic alternative to both OT and TFT, referred to as “graduated and reciprocated initiatives in tension reduction” (GRIT). GRIT, as a bargaining strategy, aims at reducing and controlling tension levels between superpowers, gradually creating “an atmosphere of mutual trust within which negotiations on critical military and political issues will have a better chance of succeeding,” and launching “a new kind of international behaviour that is appropriate in the nuclear age” (88). For the best chance of success, unilateral initiatives designed to improve relations should be “announced prior to their execution” and identified as part of a “deliberate policy of reducing and controlling tensions” (99). GRIT represents an effort to “transform a cycle of escalating tension into de-escalation.” To achieve this, “one party must first announce recognition, in general terms, of the mutual need for a changed atmosphere and then announce (and precisely perform) a sequence of specific conciliatory actions until the intent of the initiator becomes believable to the target party and reciprocation occurs” (Lindskold, Betz, and Walters 1986, 101). Although a certain level of measured retaliation is important if these cooperative initiatives are continually exploited, the initiator, nevertheless, eventually should return to a strategy of conciliation.

Lindskold, Betz, and Walters (1986) and Patchen (1987) provided additional support for GRIT. The evidence about bargaining cited by Lindskold, Betz, and Walters, for example, suggested that “an atmosphere of either cooperation or competition can be quickly altered ... when one party clearly acts inconsistently with what is characteristic of that relationship.”⁵ In an atmosphere of conflict, for example, cooperation can be restored if one party “acts deliberately – not accidentally – and clearly – not ambiguously – in a manner incompatible with that sort of relation” (1986, 99–100). Of course, the advocates of GRIT are not suggesting that such a transition would be easy;

sudden movements from conflict to cooperation are infrequent in world history. As Lindsfold, Betz, and Walters pointed out, "with a long history of mistrust, conflicting parties will tend to see manipulative intent in acts that superficially appear to be conciliatory" (101).⁶ Osgood (1962) also observed, within the U.S.-Soviet context, that the history of animosity made it more likely that, at least initially, cooperative acts would be viewed with great suspicion.⁷

However, the very complexity of international relations may actually contribute to the transition. Fisher's (1964) line of argument supports this claim: "Fractionating complex conflicts into distinct issues to be dealt with individually can be an effective way of making such conflicts manageable" and of facilitating a more cooperative relationship.⁸ Strong evidence was cited by the research team in support of GRIT, particularly from the second of two experiments. According to Lindsfold, Betz, and Walters, "The subjects quickly joined in cooperation with an adversary who abandoned a destructive strategy and announced recognition of mutual interdependence and the intention to be cooperative. The subjects virtually immediately altered their competitive pattern no matter how it was instigated – insult, attempt to coerce, challenge, or simply nonverbalized noncooperation" (1986, 101). In terms of a direct reply to provocation, therefore, GRIT implies that the response should be *more* cooperative than the provocation.⁹

With regard to a response by the United States to potential Soviet coercion, however, GRIT does not recommend abandoning a reciprocating strategy. If the Soviet Union used coercive tactics, perhaps in an attempt to test U.S. resolve, such moves were to be responded to in kind (Patchen 1988). The same argument could be made with respect to provocations directed toward the USSR by the United States. Although a precise level of retaliation relative to any given instance of provocation is not specified in the model, that information is not essential for the purpose of subjecting GRIT to empirical testing. It is clear that retaliation should be less intense than provocation.

All three of the models discussed so far (TFT, OT and GRIT) suffer from one important shortcoming: each fails to consider the "probability of errors" in pursuit of the respective strategies (Mueller 1987, 692). More specifically, the phenomenon of *noise*, namely, "any deviation from the assumption of perfect information among the players," is neglected in these strategies. If the probability of error is incorporated into decision models, "the best strategy to establish cooperation in a world of noncooperation is GRIM, the strategy that begins with a cooperative move, but never cooperates again, once a defection occurs." The logic behind GRIM's potential for success as a *competing* strategy (especially against TFT) is based on exploitation of the opponent's initial

“niceness.” Also, occasional cooperative moves (a product of noise) “inevitably result in a new exploitation before retaliation resumes” (711).¹⁰

There is a problem, however, with Mueller’s exposition, which is related to GRIM’s potential to transform conflict into cooperation: How does it create an incentive for the adversary to cooperate? It is conceivable, for example, that noise might also cause the GRIM player to misinterpret a cooperative act and accidentally (or prematurely) defect. Consequently, noise would render TFT the preferred strategy, because a noncooperative act could be misread as being cooperative, thus ending a long stream of defect/defect ($c’c’$) in a TFT game.¹¹ It could be argued, of course, that the success of GRIM is *not* a product of its ability to transform conflict into cooperation, simply because the strategy inevitably leads to $c’c’$ with any accidental or intentional defection (that is, it is extremely unforgiving). Instead, its effectiveness may be related to the *threat* of using GRIM as a bargaining strategy. In other words, although GRIM may not be an appropriate technique to create cooperation, it may prevent further escalation towards mutual defection.

Of the four alternatives, the last appears to be the accepted wisdom among realists. Coercive bargaining tactics and threats to employ force are viewed as inducements, “with the objective being that of coercing the other side to comply with one’s demands without having to resort to war” (Leng 1990, 2). Tactics and strategies that demonstrate resolve, and a propensity to take risks, are viewed as more likely to induce cooperation. Although not explicitly stated by Mueller, GRIM can be viewed as the converse of GRIT, a logical outcome of the realist point of view. For a GRIM player, cooperation must ultimately mean a compromise with asymmetric levels of concession. In terms of a direct response to provocation, for example, GRIM essentially would propose a strategy of noncooperation. Precise response levels are not specified by Mueller but, given the strategy’s emphasis on imperfect information and the margin for error, the retaliation-to-provocation ratio should be above 1.0.

Friedman (1986, 85–8) provided a further explanation for the logic of GRIM through discussion of “trigger strategies” within a game of rivalry. Each player expects moves that have resulted in an outcome on the possibility frontier (that is, efficient moves) to be continued. In the “threat game,” the choice of cooperation would occupy that role for the players, with (r_3, c_3) as the result. Players can be punished for deviating from expected patterns of behaviour. Thus, provocation triggers a change in behaviour of the adversary, with retaliation permanently replacing cooperation and proving very costly to the player who

has defected. This sequence, of course, corresponds to that of GRIM, which emphasizes coercion in the attempt to induce cooperation.

Before presenting a diagrammatic exposition of these models, I should note that all the basic strategies have variations that are not dealt with *directly* in this study. For example, George (1991, 379) discussed the strengths and weaknesses of twelve strategies designed to "change an existing status quo at the expense of an adversary." Of the twelve, five are *offensive* in nature – (1) blackmail, (2) limited probe, (3) controlled pressure, (4) *fait accompli*, and (5) slow attrition. The other seven comprise *defensive* techniques – (1) coercive diplomacy, (2) limited escalation of involvement to establish ground rules, (3) tit-for-tat reprisals without escalation, (4) accepting a test of capabilities within the restricted ground rules chosen by the opponent, (5) drawing the line, (6) conveying commitment and resolve to avoid miscalculation, and (7) time-buying actions and proposals that provide an opportunity to explore a negotiated settlement. However, each of these tactics represents more specific variations of the four general strategies.

Assessing the accuracy of GRIT, GRIM, TFT, and OT as descriptions of the behaviour of superpowers in times of crisis entails an indirect test of the offensive and defensive variations stipulated by George (1991). For instance, many of the defensive techniques include threats of force similar to GRIM, with each representing a different method of persuading an opponent to pull back or do something it otherwise would not. Similarly, strategy 7 includes elements of GRIT and OT, depending on the level of provocation and the extent to which the move deviates from (1,1).

A DIAGRAMMATIC EXPOSITION OF THE MODELS

As previously stated, each of the four strategies – TFT, OT, GRIT, and GRIM – provides an alternative approach toward transforming conflict into cooperation. These options are displayed by Figure 4.2. For ease of exposition, the notation from Brams and Kilgour (1988) will be used in describing the respective strategies. For a given case i ($i = 1 \dots n$), the horizontal axis in the figure shows the level of provocation (t_i), with the recommended level of response $q(t_i)$ along the vertical axis.¹² Although a mathematical exegesis will occur later, at this point it still is possible to convey each linkage in nontechnical terms. The TFT strategy appears as a straight line at a 45-degree angle, with retaliation precisely equal to provocation. For OT, the response is more threatening than the provocation at lower levels, but the reverse is true as t_i becomes more intense. The GRIT response always falls short

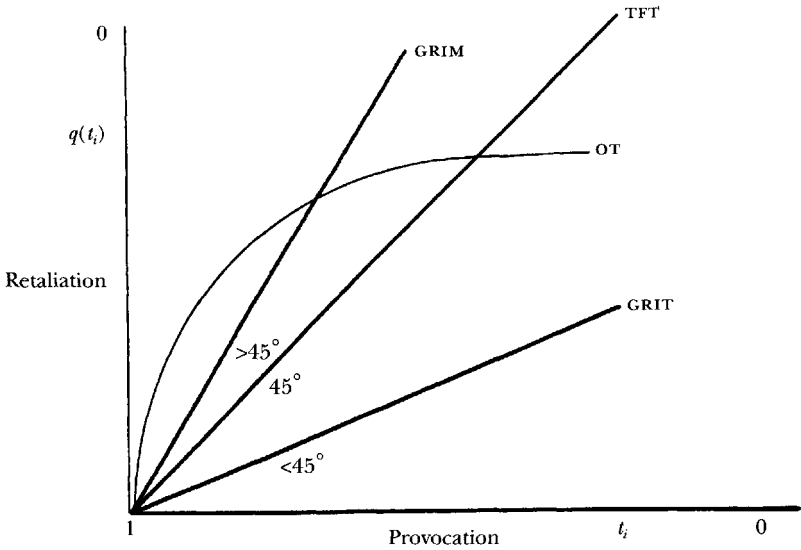


Fig. 4.2 Retaliation functions for TFT, OT GRIT, and GRIM

of the provocation level, with the gap increasing with the severity of t_i . Finally, GRIM responds more intensely than t_i at every level of provocation, also with an increasing gap at more intense levels of t_i .

Although the precise forms of the retaliation functions for GRIT and GRIM are unknown, the continuous movement away from TFT at higher levels of provocation is consistent with the logic underlying each of these models. The objective of GRIT is to control and reduce tensions. This means showing restraint, especially at high levels of tension, in order to prevent a conflict spiral. At the very least, a restrained initial response to provocation would be consistent with the overall program of reducing tensions on a global basis. As for GRIM, the higher the level of provocation, all other things being equal, the greater the need to show commitment. If the adversary has departed from cooperation in an especially dramatic way, that is evidence – from the perspective of Leng's (1990) realist – of confidence that the adversary lacks resolve.

Now that the retaliation rule for each model has been described, a basic question arises concerning U.S./Soviet crises: Which of the four models provides a more plausible assessment of the likely reaction of a superpower faced with a challenge from its rival? Before answering this question, it is appropriate to confront some of the questions underlying the application of game-theoretic models to crisis bargaining.

There is growing concern that standard applications of game theory – like those reviewed above – are counterproductive (Beer 1986; Snidal 1985; Oye 1985; Axelrod and Keohane 1985). Too many applications of game theory, Snidal (1985, 26–7) observes, “have merely been in the spirit of sorting out whether the Cuban missile crisis was really ‘chicken’ or ‘prisoner’s dilemma.’ Such usage may be helpful for reconstructing and interpreting events, but it misinterprets the primary value of game theory.” The objective is to overcome impediments to theoretical development that result from focusing on the descriptive aspects of game theory at the expense of its analytical properties. According to Snidal, game theory should be treated as a *theory* of international politics focusing on “goal seeking behaviour of states in an interdependent international system.” Some of the more central questions addressed by theorists who adopt this approach include the following:

- 1 In what ways do payoff structures determine cooperation, and to what extent can states alter situations in order to modify the payoff structure and induce cooperation?
- 2 When is cooperation, defined in terms of “conscious policy coordination,” necessary for mutual benefits?
- 3 Does iteration affect the likelihood of cooperation and, if so, what strategies can be used to lengthen the game and induce the effects of iteration?
- 4 Does the number of actors or issues affect the likelihood of cooperation?

These questions lie beyond the scope of the present investigation. An answer to the first would require “playing out” a great number of games, suggesting a more inductive approach than the one pursued here. The second question relates to the rules of the game being played and therefore is prior to the current analysis. The third and fourth questions focus on multiple stages, players and issues, thus moving beyond the analysis of superpower response to security problems in a crisis setting.

For instance, Brams and Kilgour assume that the payoff structure in the threat game (derived from chicken) never changes. Although some changes occur within the unit square (for example, movement from the preplay position to elsewhere), even these are derived from the corner payoffs, which remain unaltered. In terms of the game, the preference ordering for each player is as follows (where c' = defect, c = cooperate, and the adversary’s strategy is listed second): $c'c > cc > cc' > c'c'$. However, it may be more realistic to assume that the

payoff structure – and by implication the game itself – varies with the player's actions and events that occur outside the actor's control.¹³

Another problem concerns the use of chicken and PD as simplifications of reality. Although the models developed by Brams and Kilgour and Axelrod are “well defined, tidy, parsimonious, and orderly” depictions of social reality, “with clear definitions of rules, players, strategies, and outcomes,” international relations tends to be more complex (Oye 1985, 1–24). In terms of superpower rivalry, for example, the matrices and their respective payoff structures are not sufficient to account for one of the more important components of successful deterrence, the *credibility* of a threat. To compensate for this, the models invariably assume that each threat made in the game is perceived by the opponent as perfectly credible. Credibility, however, is a function of commitment and resolve and is not dependent on payoff structures alone. Consequently, there is very little in existing models that can accurately account for it. Of course, it may be argued that “resolve,” in and of itself, is a function of the payoff structure in a game.

However, the validity of that argument depends on *fully* accepting the assumptions associated with rational choice theories of international bargaining. Of course, as Bueno de Mesquita and Lalman point out, “If this and other assumptions help make sense of historical patterns of behaviour, *then they are useful and do not excessively simplify and trivialize what is undoubtedly a very complex process*” (1992, 27; emphasis added). In other words, whether these assumptions are useful or trivialize the analysis of international relations is an empirical question.¹⁴

FORMAL PRESENTATION OF THE MODELS

Each of the four models is presented in the context of the threat game (fig. 4.1), with one modification to it. James and Harvey (1989) note that in actual crises simultaneous noncooperation is unknown. For ease of exposition, it will be assumed that Row is initially the more cooperative player (that is, $s = 1$), with Column committing the act that initiates the crisis. The time lag involving t_i and $q(t_i)$, however, is not long enough to introduce complications such as discounted future payoffs.

It might be suggested that this approach will generate an inappropriate measurement of the performance of TFT because TFT is expected to find favour in an iterated game of PD. With the shadow of the future playing a prominent role, convergence toward TFT is expected to occur over a sequence of play. However, in a superpower crisis, possibly at the brink of nuclear war, the long-term viability of the relationship is a secondary matter. If superpower response to provocation does not approximate TFT, then its relevance to the playing out of

intense international conflicts – which tend to involve relatively quick decisions – may be limited. A static, or single-play, testing environment constitutes the greatest challenge for TFT as a potential description of strategic choice. Of course, the possibility of nuclear war as the result of a sequence of mutual defections still would create significant interest in restoring cooperation. Patchen has argued that a TFT strategy would be more effective in a chicken situation “since an adversary suffers his worst outcome when his competitive move is reciprocated in chicken, as compared to suffering his next-to-worst outcome when this occurs in prisoner’s dilemma” (1988, 280–1).

Points of clarification arise for the other strategies as well. For GRIT, gradual change plays a central role (Osgood 1962). Its logic implies that an initial provocation – such as that which marks the outset of a crisis – should be met with a less coercive, as opposed to matching, response. Thus evidence of GRIT should be available even in a single-play environment. The same is true for GRIM. In fact, *all* the strategies face the strictest possible testing environment. Learning and convergent expectations cannot be expected to affect the results of an isolated interaction; underlying conformity to any one of the models would be most weakly revealed under these circumstances. In other words, any evidence of regularity in responses to provocation should be seen as extremely favourable to the strategy in question.

For a given case, the margin of error for TFT corresponds to the difference between the levels of provocation by one player and retaliation by the other. Once again, in order to simplify the notation, in all instances it will be assumed that Column is the initiator and Row the target. Equation 1 represents the average level of the discrepancies, F_1 . Squared differences will be summated to create F_2 , an alternative measurement of the margin of error. The advantage of using F_2 is that it is more sensitive to the impact of larger errors. When comparing TFT to OT, the two versions of the margin of error may reveal some differences.

$$(1) F_1 = 1/n \sum_{i=1}^n |t_i - q(t_i)|,$$

where

F_1 = the average difference between predicted and observed responses;

n = number of cases ($i = 1 \dots n$);

t_i = provocation by the initiator (also the predicted response by the defender) in case i ;

$q(t_i)$ = the observed response by the target in case i .

Brams and Kilgour (1988) have developed a more complex rule for optimal response. The retaliation function $q_1(t_i)$ depends on the level of provocation (t_i) and the relative magnitudes of the corner payoffs. Payoffs at the final position are weighted averages of the "pure" payoffs at the corners of the game board. The payoff functions are bilinear, meaning linear in each coordinate. In calculating payoffs for Row and Column at any point (x,y) , referred to as $P_R(x,y)$ and $P_C(x,y)$, respectively, the payoffs at each corner are "weighted by the product of the distances, parallel to the axes, from (x,y) to the opposite corner." The payoffs $P_R(x,y)$ and $P_C(x,y)$ are therefore calculated as follows (Brams and Kilgour 1987a, 837):

$$\begin{aligned} (2) \quad P_R(x,y) &= (1-x)(1-y)r_1 + (1-x)yr_2 + xy r_3 + x(1-y)r_4 \\ (3) \quad P_C(x,y) &= (1-x)(1-y)c_1 + (1-y)xc_2 + xy c_3 + y(1-x)c_4. \end{aligned}$$

Although another weighting system would alter the proportions of the resulting payoffs, as Brams and Kilgour observe, it would not alter "the basic nature" of the Nash equilibria within the game (1988, 107, 108).

The level of retaliation that is just sufficient to deter the adversary from further aggression is given by equation 4:

$$(4) \quad q_1(t_i) = \frac{c_3 - t_i c_2}{1 - t_i(1 - c_3 + c_2)}.$$

With $t_i = 1$, the function recommends continued cooperation. When Column engages in provocation, meaning $t_i < 1$, retaliation is called for at some level.

Having identified the retaliation function, it is possible to specify the margin of predictive error for OT:

$$(5) \quad E_{j,k} = 1/n \sum_{i=1}^n \left| \frac{(c_3 - t_i c_2)}{1 - t_i(1 - c_3 + c_2)} - q(t_i) \right|,$$

where $E_{j,k}$ = the average (unsquared) difference between predicted and observed response under payoff configuration j ($j = 1 \dots 36$). For all $E_{j,k}$, $k = 1$. Although the values of c_2 and c_3 are not specified by Brams and Kilgour, the boundaries are $0 < c_2 < c_3 < 1$. The limits of 0 and 1 represent the fixed values of c_1 and c_4 , representing nuclear war and victory, respectively. Squared deviations will also be used, with $E_{j,k}$ having 36 versions and $k = 2$ in all cases, to distinguish these values from the unsquared series, denoted by $k = 1$.

Table 4.1 shows 36 value combinations (V_i , $i = 1 \dots 36$) for the intermediate payoffs, c_2 and c_3 . Each of the 36 pairs has the basic property $c_3 > c_2$, with the size of that difference being allowed to vary

Table 4.1
Intermediate Payoff Combinations and Resulting Error Levels

Payoff Combination V_i ($i = 1 \dots 36$)	Payoff Values		Average Error Level for Unsquared and Squared Differences	
	c_2	c_3	$E_{j,k} \ j = 1 \dots 36;$ $k = 1$	$E_{j,k} \ j = 1 \dots 36;$ $k = 2$
V_1	0.8	0.9	0.20	0.07
V_2	0.7	0.9	0.21	0.07
V_3	0.6	0.9	0.22	0.08
V_4	0.5	0.9	0.22	0.08
V_5	0.4	0.9	0.23	0.08
V_6	0.3	0.9	0.23	0.08
V_7	0.2	0.9	0.24	0.08
V_8	0.1	0.9	0.24	0.08
V_9	0.7	0.8	0.15	0.04
V_{10}	0.6	0.8	0.16	0.05
V_{11}	0.5	0.8	0.17	0.05
V_{12}	0.4	0.8	0.18	0.06
V_{13}	0.3	0.8	0.18	0.06
V_{14}	0.2	0.8	0.19	0.06
V_{15}	0.1	0.8	0.20	0.07
V_{16}	0.6	0.7	0.12	0.03
V_{17}	0.5	0.7	0.14	0.04
V_{18}	0.4	0.7	0.15	0.04
V_{19}	0.3	0.7	0.15	0.05
V_{20}	0.2	0.7	0.16	0.05
V_{21}	0.1	0.7	0.17	0.05
V_{22}	0.5	0.6	0.14	0.03
V_{23}	0.4	0.6	0.13	0.03
V_{24}	0.3	0.6	0.13	0.04
V_{25}	0.2	0.6	0.14	0.04
V_{26}	0.1	0.6	0.15	0.04
V_{27}	0.4	0.5	0.20	0.05
V_{28}	0.3	0.5	0.17	0.04
V_{29}	0.2	0.5	0.15	0.04
V_{30}	0.1	0.5	0.14	0.04
V_{31}	0.3	0.4	0.27	0.08
V_{32}	0.2	0.4	0.22	0.06
V_{33}	0.1	0.4	0.19	0.05
V_{34}	0.2	0.3	0.34	0.13
V_{35}	0.1	0.3	0.28	0.09
V_{36}	0.1	0.2	0.40	0.19

at intervals of 0.1. The absolute magnitudes of c_3 and c_2 range from 0.9 to 0.2 and 0.8 to 0.1, respectively. Each payoff combination will be used to generate a separate value ($E_{j,1}$, $j = 1 \dots 36$) for the margin of predictive error specified by equation 5. A parallel set of values ($E_{j,2}$, $j = 1 \dots 36$) will be created using the squared differences. The reason for this elaborate scheme of measurement is that, unlike the other models, the margin of error for OT depends on the values of the payoffs. If, for example, all 36 versions of $E_{j,1}$ were lower (higher) in magnitude than F_1 , that would constitute very strong support (refutation) for OT as compared to TFT. By contrast, with only one measurement scheme, much less could be said about relative performance by OT, given the lack of a priori reasoning in favour of a specific pair of values for c_2 and c_3 .¹⁵

One further point of explanation is in order with respect to the values of pure payoffs. It is obvious intuitively that, on a scale ranging from 0 to 1, the boundaries must be nuclear devastation and victory. But what is the substantive meaning of the rest of the scale? It probably is impossible to estimate the genuine values of compromise and defeat relative to each other – or the boundaries – on a consistent basis. In fact, the specific utilities might change even within the lifetime of a given crisis. However, it is feasible to assume the relative values are fairly steady if the *underlying Soviet-U.S. rivalry* is used to assess the outcomes. Thus the experimental approach reflected by the scenarios listed in Table 4.1 can be justified as an initial effort to measure the pure payoffs associated with defeat or compromise at any given stage of the rivalry.

With regard to GRIT, there is no formal retaliation function available in the literature. However, it is possible to infer two properties involving the relative magnitudes of provocation and retaliation. First, since reduction in tension is the underlying objective of the response, it is expected that $q(t_i)$ will be greater in magnitude (that is, *less* coercive) than t_i . Second, there should be a negative correlation between t_i and $(q(t_i) - t_i)$. As t_i becomes more intense, the dampening of retaliation should become greater. In other words, as t_i approaches 0, the magnitude of the discrepancy between retaliation and provocation should increase.

GRIM also lacks a fully specified retaliation function. However, it can be inferred that the linkages expected for GRIT should be reversed for GRIM. First, it is expected that $q(t_i)$ will be lower (that is, more coercive) than t_i , given concerns about probability of error and imperfect information. Second, there should be a positive correlation between t_i and $(q(t_i) - t_i)$. With a more intense level of provocation (that is, t_i approaching 0), the relative severity of retaliation is expected to become greater.

Another complication that might be raised is that of the unitary actor assumption (Allison 1971): Does a single preference ordering exist for either the United States or the USSR, given the range of actors involved on each side in the formulation and implementation of foreign policy? Among others, Bueno de Mesquita (1981) and McGinnis (1988) have argued that in crisis situations (and especially those involving long-term rivals), decision making tends to converge toward *de facto* unitary action. Time constraints, in particular, make it less likely that actors outside the executive's inner circle will play a role in developing and carrying out policy. This assumption is even more likely to be justified in strategic nuclear interactions, where time limitations are most severe.

Measuring Provocation and Retaliation Levels⁶

Based on criteria described in detail in chapter 3, a subset of twenty-eight cases of superpower rivalry from the ICB dataset on actor-level crises has been selected.¹⁷ As noted, with regard to the time frame of the analysis, only crises in progress in 1949 or thereafter are appropriate for inclusion. In practical terms, this entails the possibility, however remote, of a nuclear exchange, producing (r, c_1) from Figure 4.1. The USSR detonated a nuclear device in 1949, meaning that each superpower then had some chance of experiencing a nuclear attack – even if limited and through a client state – at some inexact point during that year. Although any boundary would be arbitrary, the choice of 1949 excludes all crises for which this could not happen as intended. It also includes every crisis in which that outcome had some chance of occurring. The list in Table 3.1 ends with crises initiated in 1985, the current terminal point of the ICB project's data set.¹⁸

With the cases having been identified, measurements for t_i and $q(t_i)$ are required before testing the models. The scale values for the polar points of cooperation and noncooperation are 1 and 0, respectively. For intermediate points, values are generated using an adapted version of the measurement scheme developed by James and Harvey (1989).¹⁹ The basic property of that scale is an exponential decrease in values; in other words, as the intensity of a provocation or retaliation increases, its score approaches 0 – meaning nuclear attack – more quickly. Table 4.2 displays the 15-point scale of superpower bargaining techniques.²⁰ The points on the scale combine the coding of the relevant ICB variables, “triggering action” and “major response” (Wilkenfeld, Brecher, et al. 1988). The scale points are generated by the following expression:

$$(6) \quad s'_k = 2 - e^{0.693(s_k)},$$

Table 4.2
Superpower Bargaining Techniques: Provocation and Retaliation

<i>Intensity</i>	<i>Score</i>
1 Nuclear strike	0.000
2 Full-scale conventional war	0.098
3 Indirect full-scale conventional war	0.198
4 Serious clashes	0.275
5 Indirect serious clashes	0.360
6 Minor clashes	0.438
7 Indirect minor clashes	0.514
8 Multiple techniques, including violent military	0.585
9 Nonviolent military techniques	0.654
10 Multiple techniques, including nonviolent military	0.720
11 Economic act	0.781
12 Political act	0.841
13 Verbal act	0.896
14 No action	0.950
15 Cooperation	1.000

where s'_k = the transformed (exponential) scale point ($k = 1 \dots 15$), and s_k = the linear scale point. Each s_k has the value $(k - 1)/14$, with $s_1 = 0$, $s_2 = 1/14 \dots s_{15} = 1$. The coefficient for s'_k , 0.693, is appropriate in the exponent, given the boundaries intended for s'_k . When $s_k = 0$, the transformation results in $s'_k = 1$; with $s_k = 1$, $s'_k = 0$. The scale in Table 4.2 covers the full range of potential values for cooperation and noncooperation.²¹ It has a fixed zero point, nuclear strike, which represents the absence of cooperation. At the other extreme lies cooperation itself.²² Although the largest interval on the scale is the one separating a nuclear strike from full-scale conventional war, no *qualitative* difference in the range of values is assumed. The identification of a zone of uncontrollable escalation in Figure 4.1 would justify the incorporation of a step-level in the scale, with the penultimate point falling substantially below a nuclear strike. Fortunately, from a practical point of view the level of coercion sufficient to bring about irreversible escalation is unknown; a continuous range of values therefore is appropriate.

Values for t_i and $q(t_i)$ are displayed for each case from Table 4.1 in the final two columns of Table 4.3.²³ A graphic illustration of the provocation and retaliation scores for each case appears in Figure 4.3.

Although the advantages of using the ICB cases in testing the four models have already been discussed, a few more specific, measurement-related benefits should be noted. First, every foreign policy crisis has an explicitly designated triggering act and major response, making it possible to identify circumstances in which one superpower caused

Table 4.3
Trigger and Response in Foreign Policy Crises, 1948–88

<i>Case</i>	<i>Provocation by</i>	<i>Retaliation by</i>	<i>Trigger t_i</i>	<i>Response $q(t_i)$</i>
1 Berlin blockade	USSR	USA	.65	.72
2 Berlin blockade	USA	USSR	.84	.65
3 Korean War I	USA	PRC/USSR	.59	.72
4 Korean War II	USSR/PRC	USA	.59	.65
5 Korean War II	USA	USSR	.19	.65
6 Korean War III	USA	PRC/USSR	.65	.84
7 Taiwan Straits I	USA/Taiwan	PRC/USSR	.84	.36
8 Taiwan Straits I	USA/Taiwan	PRC/USSR	.84	.36
9 Suez-Sinai Campaign	USSR	USA/Britain/France	.65	.72
10 Syria-Turkey	USA/Turkey	USSR/Syria	.72	.90
11 Taiwan Straits II	USA/Taiwan	PRC/USSR	.65	.90
12 Berlin deadline	USSR	USA/Britain/France	.84	.84
13 Berlin deadline	USA/Britain/France	USSR	.65	.72
14 Bay of Pigs	USA	USSR/Cuba	.36	.90
15 Berlin Wall	USSR/GDR	USA/FRG	.72	.65
16 Cuban missiles	USSR/Cuba	USA	.65	.65
17 Cuban missiles	USA	USSR/Cuba	.65	.72
18 Congo I	USA/Belgium	USSR	.28	.90
19 Gulf of Tonkin	USA	USSR/North Vietnam	.28	.72
20 Six Day War	USSR	USA/Israel	.72	.65
21 War of Attrition	USSR/Egypt	USA/Israel	.65	.72
22 Cienfuegos Bay	USSR	USA	.65	.90
23 Yom Kippur War	USSR/Egypt/Syria	USA/Israel	.72	.72
24 Ogadan III	USA	USSR/Ethiopia	.65	.84
25 Invasion of Grenada I	USA	USSR/Grenada	.28	.84
26 Invasion of Grenada II	USSR/Grenada	USA	.65	.28
27 Nicaragua, Mig 21s	USA	USSR/Nicaragua	.72	.90
28 Nicaragua, Mig 21s	USSR/Nicaragua	USA	.65	.72

the other to experience and cope with crisis conditions. Second, the ICB data set contains the only exhaustive record of superpower crisis behaviour over the relevant time period. Third, data sets that focus on war or other types of conflict are less suitable for testing models of response to provocation because the behaviour of the superpowers is not always distinguished on an individual basis within the coding. Fourth and most essential, the ICB project includes a case history, with an attendant bibliography, for every case (Brecher, Wilkenfeld, et al. 1988; Wilkenfeld, Brecher, et al. 1988). This makes it possible to

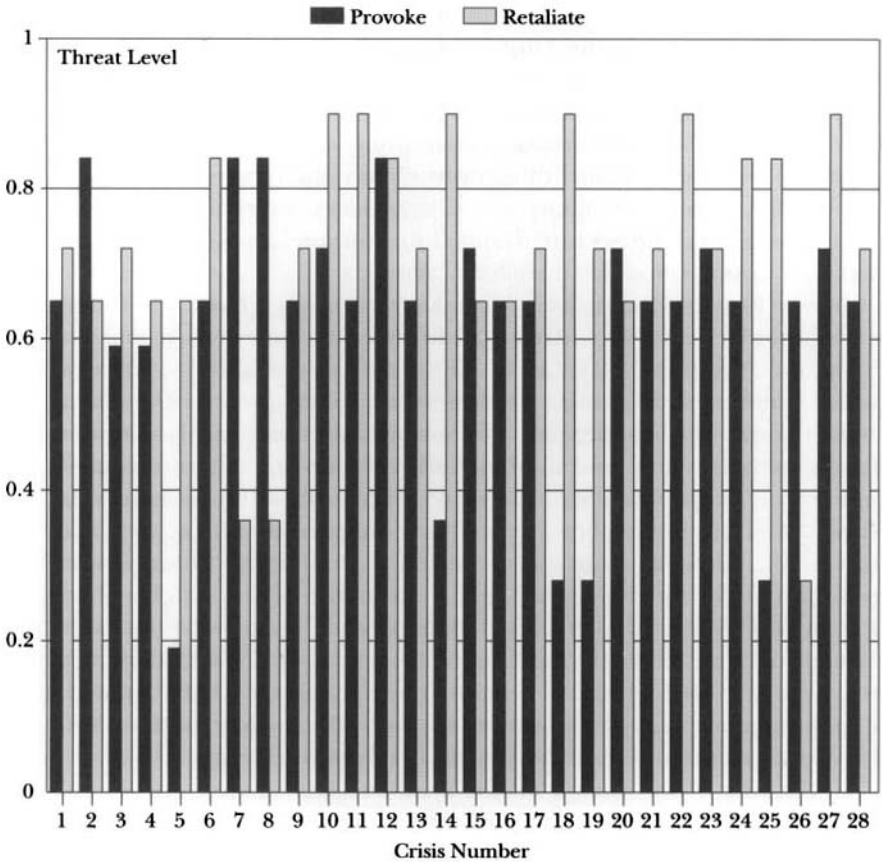


Fig. 4.3 Provocation and retaliation in superpower crises

probe more deeply for potential inaccuracies before making a final decision about selection of superpower rivalry cases and subsequent coding.

A description of a few cases will help clarify the coding procedures used for Figure 4.3. In the case of Korean War I, the actions taken by the United States on 27 June 1950 included the authorization by President Truman for U.S. forces to fight alongside Republic of Korea troops (a violent military act), nonviolent military manoeuvres by the Seventh Fleet in the Formosa Straights, and a series of political actions in the form of U.S.-sponsored resolutions in the United Nations Security Council. The overall level of provocation by the United States that triggered a crisis for the USSR, therefore, was coded (Table 4.2) as .59 – multiple, including nonviolent military, actions. The *response* by the

People's Republic of China (PRC) was simply a speech by Premier Chou En-Lai accusing the United States of aggression (.896), while the USSR provided Shanghai and other cities with air cover in response to possible bombing from Taiwan (.65). Given this particular sequence of threats and counterthreats, the response by the USSR/PRC was coded as .72 (multiple, including nonviolent military, acts) – examples are Chou En-Lai's accusations of U.S. aggression, a verbal act; absence of the USSR from the Security Council for the vote, a political act; and USSR air cover, a nonviolent military action.

In one Korean War II case, South Korean forces crossed the thirty-eighth parallel on 1 October 1950, followed by American troops on 7 October. From the Soviet and Chinese perspective, the U.S. advance was perceived as an indirect escalation to full-scale war with a client state (.198). The response by the Soviet Union was to place Soviet forces in the Far East and in Manchuria on alert status (a nonviolent military act, .65). The second foreign policy crisis within Korean War II occurred on 31 October 1950, when confirmation of the presence of Chinese forces in Korea was received in Washington, triggering a crisis for the United States. The U.S. response was a statement by President Truman that the United States would use any weapons in its arsenal needed to defeat communist forces (.65). Clearly Truman's response was more provocative than a mere verbal act (.896). Indeed, it was tantamount to a declaration of war, with nuclear undertones, further escalating the crisis to potentially disastrous levels.

In the case of the Cuban missile crisis of 1962, there were several responses to the blockade, a nonviolent military action (.65) that triggered a crisis for the Soviets. In addition to accelerating their deployment plans (a nonviolent military response), the Soviets repeatedly denied the existence of the missiles in the United Nations (a political act), sent two letters to Washington (a verbal act), and ultimately bargained their way out of the crisis through a compromise over the missiles in Turkey. Given these actions, the overall response was coded as multiple, including nonviolent military actions (.72).

Finally, the foreign policy crisis experienced by the United States over the Cienfuegos Bay base in Cuba started on 16 September 1970.²⁴ On that date, National Security Advisor Henry Kissinger informed President Nixon that U.S. intelligence flights had substantiated reports about Soviet construction of a nuclear submarine base in Cuba. This activity triggered a crisis for the central decision makers of the United States, who perceived a threat to the 1962 agreement on Cuba. The United States viewed nuclear missiles stationed off the coast of Florida as a serious danger to military security. Since the base might soon become operational and capable of offensive action, the United States

had a limited time available for response. The nonviolent military trigger by the Soviets results in $t_i = 0.65$. The United States responded on 25 September – the day that the story broke in the U.S. press – with a demand by Kissinger for an explanation. This riposte is coded as a verbal act, so $q(t_i) = 0.90$. The crisis did not escalate further. It ended on 23 October when Washington received Soviet assurances “that construction had been halted and that the Soviet naval force had left Cienfuegos.”

ANALYSIS OF DATA

Comparison of the predictive error for TFT and OT is favourable to the latter. Using equation 1 (above), $F_1 = 0.22$. Among the 36 versions of $E_{j,1}$ generated by equation 5, 25 (69.4%) of those listed in Table 4.1 fall below that level (see Figures 4.4 and 4.5). Based on the squared differences, $F_2 = 0.081$. Among the 36 versions of $E_{j,2}$ listed in the table, 28 (77.2%) fall below that level. Although not overwhelming, these results favour OT over TFT. The squared differences are more favourable to OT, suggesting that TFT tends to make more large errors, although in some of the scenarios these are more than balanced out by very close margins on other cases.

It is interesting to note that among the combinations favourable to TFT, six of the eight with $c_3 = 0.9$ are included. Compromise is virtually equated with victory when the magnitude of c_3 is so great. A natural question is to ask whether such configurations of payoffs for the pure outcomes are realistic, by and large, in superpower crises. Since any crisis entails some danger of escalation to nuclear war, a negotiated outcome is potentially attractive. A reasonable conjecture is that the more fundamental the issue, the less likely it is that compromise will be regarded as nearly equivalent to victory. However, the relative importance attached to the values threatened certainly will affect the evaluation of c_3 . The investigation in chapter 5, in part, will assess the impact on strategic choice of the issues under dispute in a specific case. The assumption of near equivalence for compromise and victory might be expected to hold within some range of cases, while for others a greater discrepancy would be anticipated. Perhaps TFT is a more appealing strategic option in the first environment, whereas OT finds favour in the second.

At the other end of the spectrum are $E_{16,1} = 0.12$ and $E_{16,2} = 0.03$, the smallest error terms among each of the two sets of 36 versions of OT. These minimal values are generated when $c_3 = 0.7$ and $c_2 = 0.6$, meaning that compromise and defeat are viewed as very similar outcomes. It is especially interesting to consider whether or not the

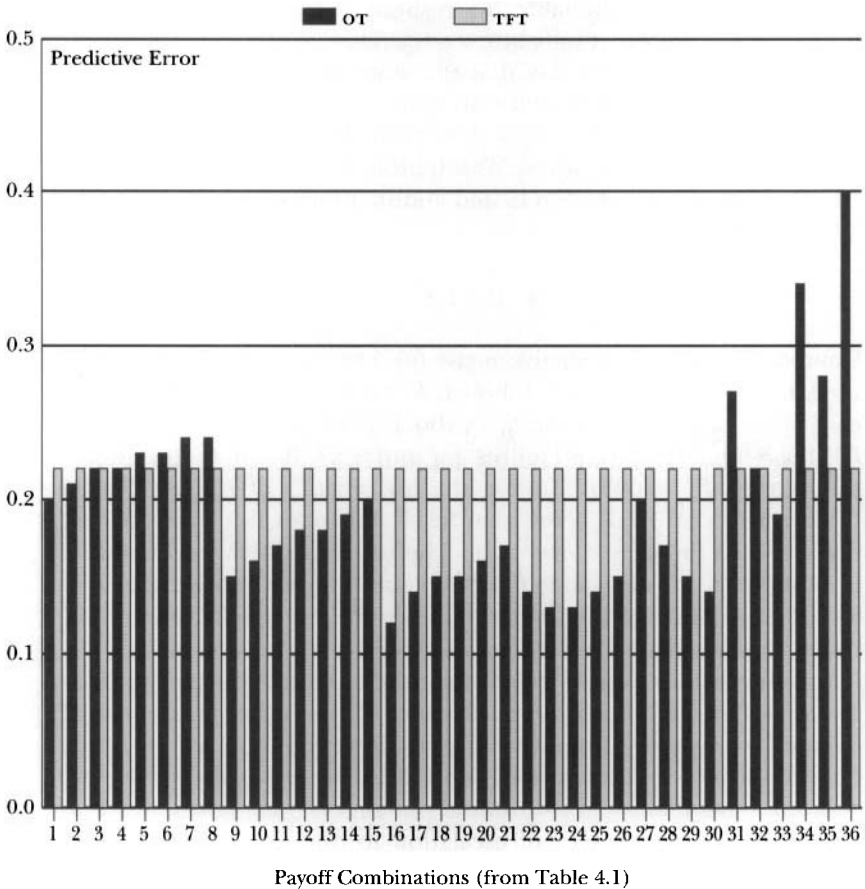


Fig. 4.4 Margin of predictive error, OT vs TFT

payoffs just noted would be accurate for a wide range of cases. If so, that would constitute overwhelming support for OT as compared to TFT. The closely matched values for defeat and compromise, intuitively speaking, suggest that the issue(s) under dispute are especially significant. Each outcome is somewhat removed from victory, or full achievement of goals (1.0 versus 0.6 or 0.7). But, given the catastrophic nature of mutual escalation, there is a sense of realism in the fact that each of the two intermediate payoffs is even further away from 0. Given the preceding analysis of the configurations with $c_3 = 0.9$, which favoured TFT, the notion of crucial cases leading greater support to OT becomes even more intriguing. Once again, it would be interesting to explore the role of the centrality of value threats in determining strategic choice.

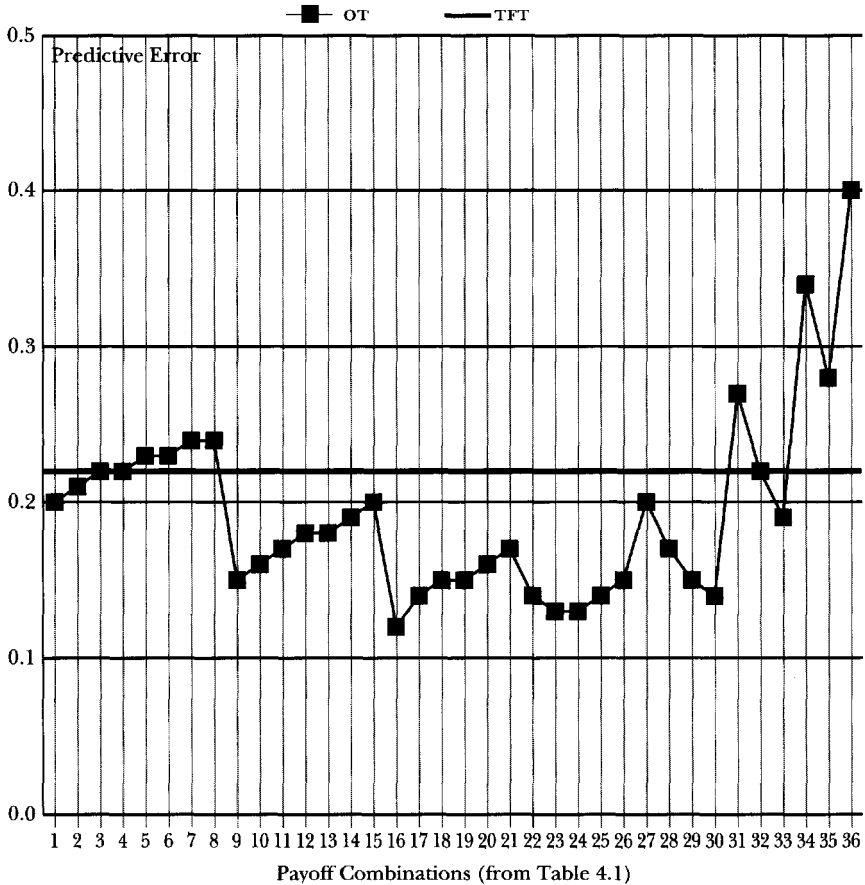


Fig. 4.5 Margin of predictive error, OT vs TFT

While GRIT and GRIM have not been formally specified as models, there are expectations regarding the relative magnitudes of t_i and $q(t_i)$. There are three cases in which the levels are equal, nineteen with $t_i < q(t_i)$ and six with the reverse ordering. This is somewhat favourable to GRIT, although not overwhelming. The cases in which the retaliation is more intense than the provocation are the Berlin blockade (the first of its two cases), both crises related to Taiwan Straits I, the Berlin Wall, the Six Day War, and the invasion of Grenada. These cases are diverse and do not include the full subset of extremely intense crises from the Cold War; among those missing are the Korean War sequence, one crisis over the Berlin blockade, and the Cuban missile crisis. It is also interesting to note that, among the six cases, the provocation is issued by the USSR on three occasions, with the

United States in that role in three others. Thus, it is inappropriate to fall back on the claim that, even if GRIM appears to be less common, one of the superpowers follows its dictates more than the other.

More decisive evidence in favour of GRIT over GRIM is supplied by the correlation of $(q(t_i) - t_i)$ with t_i ($r = -0.7633$, $p < 0.001$, $N = 28$). As the intensity of the provocation increases (that is, as t_i approaches 0), the *relative* severity of the retaliation decreases, that is, $(q(t_i) - t_i)$ becomes larger. It is interesting to note that intense provocations are especially unlikely to generate an equal or more intense response. Consider the initiatives that match or exceed the violent military level, that is, where t_i is less than or equal to 0.585. In all five of these instances, the retaliation is less coercive than the provocation, which seems especially out of line with GRIM's realist perspective. However, it is consistent with the realist dimensions of OT, suggesting that, at least in some cases, a more than proportional response is useful.

A logical next step for the current program of research is to see whether approximation to OT (or any other strategy) can be linked to crisis outcomes. Presumably, selection of a given level of retaliation is expected to have implications for the degree of satisfaction ultimately achieved by the initially more cooperative player (hereafter referred to as the "defender"). If so, that would be of considerable practical value, going beyond the description of strategic choice and into its consequences for foreign policy. As Snyder and Diesing point out, "Whether to be firm and tough toward an adversary, in order to deter him, but at the risk of provoking his anger or fear and heightened conflict, or to conciliate him in the hope of reducing sources of conflict, but at the risk of strengthening him and causing him to miscalculate one's own resolve, is a perennial and central dilemma of international relations" (1977, 254). Proponents of each strategy assume that outcomes will be favourable to the state pursuing the respective response pattern. GRIT, for instance, recommends conciliation and a more cooperative approach to crisis management. As noted in chapter 1, advocates of the "reassurance" school claim that GRIT would be the most effective crisis management technique because it creates an environment of mutual trust wherein mistakes, misperception, and other decision pathologies can be avoided (Lebow and Stein 1989a, b). In contrast, GRIM proposes a strategy of coercive diplomacy and noncooperation – deterrence rather than reassurance. Phrases like "competition in risk taking," "exploitation of potential force" (Schelling 1960, 69–78), and "diplomacy of violence" (Schelling 1966, 3) are often used to characterize the underlying logic and rationale of GRIM. The question is whether outcomes in superpower disputes are consistent with expectations of success underlying each strategy.

Studies by Leng (1988) provide useful examples of this type of research. Leng focused on the best *opening* strategy in a crisis and found that none of the nuclear games (that is, those with the structure of Chicken) confers an advantage upon a player "who begins with a more coercive bargaining strategy than it employed in the previous crisis" (188). These results confirmed earlier findings from a study of influence strategies in serious international disputes by Leng and Wheeler: "if our results suggest anything, it is that an assertive, bullying strategy is both less effective and more risky than much of the folklore of power-based policies would have it" (1979, 681).

Leng (1990) examined the behaviour of states in forty crises between 1816 and 1980 using Correlates of War data. The objective was to determine which of the two approaches ("realism" or "reciprocity") most accurately predicted how states responded to a variety of attempts at influence during militarized crises. Leng found that the reciprocity model was "clearly and consistently" superior to the realist model as a description of state behaviour. "The prescriptive implications," Leng concluded, "are obvious: If governments are most likely to respond to inducements in kind, then there is little to be said for employing increasingly coercive inducements, and a great deal to be said for more accommodative influence techniques" (22).

In contrast, Corson's (1970) analysis of the Berlin Wall and Cuban missile crises supported GRIM. Corson found that higher degrees of compliance by the United States and USSR were more likely when demands were accompanied by highly intense conflictual deeds, a high ratio of coercive acts to conflictive words, and extensive use of military resources. Coercion and intimidation, in other words, seemed to work.

Given the inconsistent findings about the results of crises, further analysis is warranted. Data are available on several outcome variables for the twenty-eight cases listed in Table 3.1. Each case includes information on *goal achievement* (that is, objective assessments of victory, stalemate, and defeat) and general *satisfaction* with the outcome as perceived by the crisis actors.²⁵ These "objective" and "subjective" estimates of outcome were used to test propositions about the effectiveness of each strategy. If consistent patterns were identified for both measures, that would imply stronger support for the strategy in question.

With these considerations in mind, the predictive power of each model will be assessed in turn, starting with OT. According to Brams and Kilgour (1988), if the initially more cooperative player in a crisis retaliates at or beyond the threshold required to deter an opponent, the result should be a higher degree of satisfaction for that player. The level of retaliation just sufficient to deter aggression was stipulated

in the threat game (see equation 4, above). It is important to note, prior to assessing OT, that "symmetry" is implicit in the threat game. In a given crisis, c_4 and r_4 refer to unilateral concessions made by the adversary, with c_2 and r_2 meaning the reverse. The status quo, (c_3, r_3) , is an intermediate value reflecting the situation at the outset of the crisis and, of course, (c_1, r_1) represents nuclear devastation.

It is not clear, however, that the United States and the Soviet Union, the countries that would play the game in reality, attach equal values to the four pure payoffs. Symmetric ordering of the alternatives might be preserved – as Brams and Kilgour require – without *cardinal symmetry* holding true. The most salient potential difference concerns the pre-play position (r_3, c_3) . Perhaps the leader and challenger, to use the terminology of Organski and Kugler (1980), will view that situation differently. Compared to the USSR, the United States might be expected to see a smaller difference between c_3 and c_4 and a greater gap from c_3 to c_2 because it finds the status quo relatively more appealing. In other words, it would follow a risk-averse decision rule by seeking more intensely to preserve the status quo and avert losses. By contrast, as a challenger for global leadership, the USSR would not be inclined toward a risk-averse outlook regarding the pure states of the game.²⁶

Given these potential differences between the superpowers, measurement of payoffs in this study will allow for symmetric but unequal (where r_i is not always equal to c_i) values.²⁷ This was accomplished by including a payoff function with equal increments for the USSR (0, 1/3, 2/3, 1), while for the United States a higher relative value was placed on the status quo (0, 1/4, 3/4, 1). This scheme of payoffs will be used in testing OT once again.²⁸

Figure 4.6 illustrates the relationship between OT and crisis outcome. Of the 19 cases in which the defender threatened retaliation at or above the threshold for deterrence – stipulated by the threat game – the defender was satisfied in 13 (68.4 percent; $\tau = .32$; $p < 0.10$; $N = 28$).²⁹ Similarly, in 73.7 percent of those cases (14/19) the defender succeeded in achieving central objectives in the dispute ($\tau = .34$; $p < 0.10$; $N = 28$). By contrast, in the 9 cases where the defender did not convey a sufficient retaliatory threat in response to the adversary's aggression, the defender (that is, the initially more cooperative player) was fully satisfied in only one. As an extension of this line of argument, it was expected that approximation to OT would result in satisfaction for each side in the dispute, because it would stabilize the crisis and prevent an escalation of hostilities. This proposition was moderately supported by the data.

A final proposition consistent with OT as a model of superpower rivalry focuses on the effects of distance between the preplay position

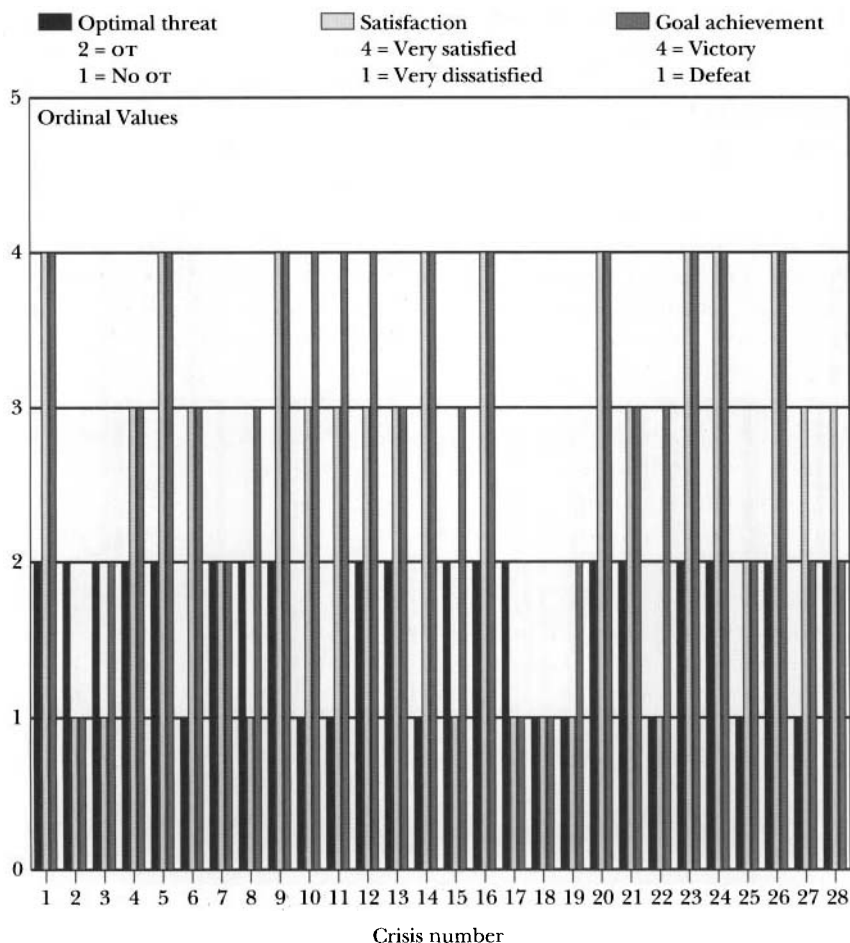


Fig. 4.6 Optimal threats and satisfaction

(1,1) and the initial position $(t,1)$. According to this proposition, a shorter distance from the preplay position to the initial position, which facilitates efforts towards crisis stability, should produce a higher level of satisfaction for the defender. As Brams and Kilgour observe, “crisis stabilization is enhanced when the advantages of mutual cooperation in the threat game are substantial for the players and the status quo at $(t,1)$ is not too far from $(1,1)$ ” (1987, 843). Since both superpowers are likely to benefit from crisis stability, prospects are better for the eventual satisfaction of the more cooperative player. This proposition was confirmed by the data. Satisfaction for the initially more cooperative player is more likely to result when the initial position is closer

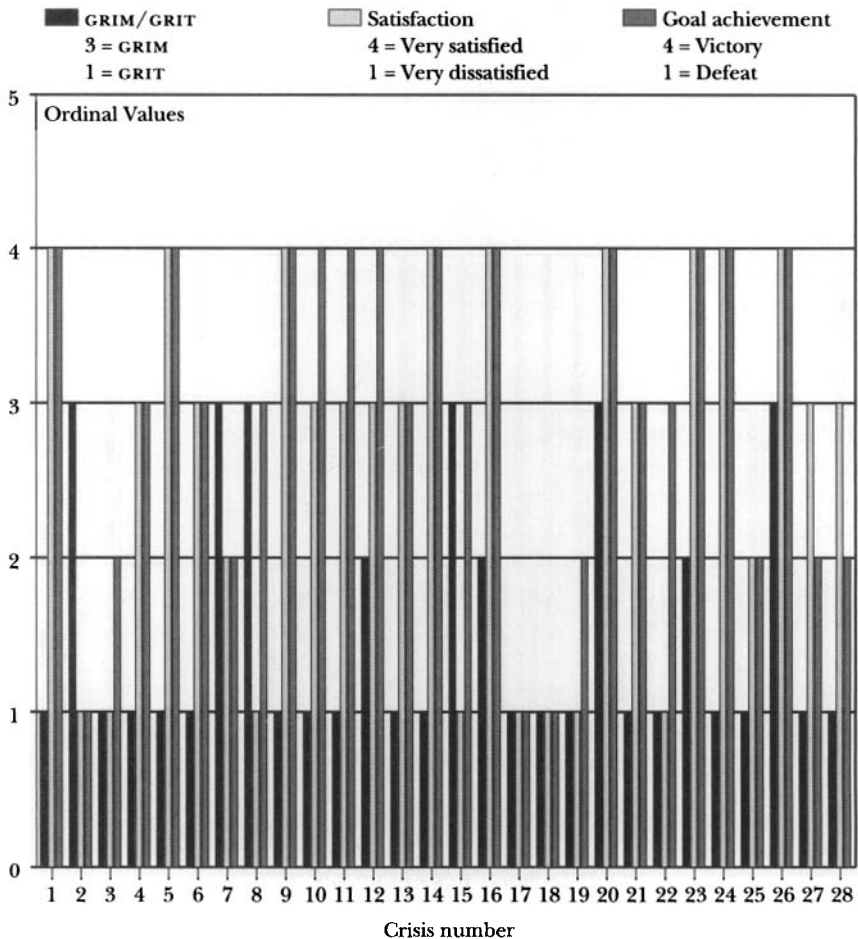


Fig. 4.7 Satisfaction and goal achievement, GRIM vs GRIT

to the preplay position ($\tau b = .31$; $p < .10$; $N = 28$). The greater likelihood of satisfaction with the outcome is consistent with intuition; when the retaliating superpower is at a position close to the preplay setting, the range of sufficiently assertive responses, or the room for manoeuvre, is increased, by definition.

With respect to comparing the effectiveness of GRIT and GRIM, Figure 4.7 displays the relationship between satisfaction and crises in which the retaliation was more threatening, equal to, and less threatening than the provocation. Of the 19 cases in which GRIT was practised, the defender was satisfied with the outcome in 13 and successfully achieved the central objectives in the dispute in 12. In

4 of the 6 crises in which GRIM was practised, the initially more cooperative player either was dissatisfied or very dissatisfied with the outcome. Overall, the results support GRIT over GRIM and, by implication, the findings of Leng and Wheeler (1979), Lindsfold, Betz, and Walters (1986), Patchen (1987), Leng (1988, 1990), Goldstein and Freeman (1990), and Goldstein (1991). Reciprocating strategies generally worked better in preventing serious militarized disputes from escalating to war. In other words, GRIT appears to be a more effective and successful crisis management technique, particularly if long-term cooperation is the ultimate goal (Goldstein and Freeman 1990).

SUMMARY

With regard to furthering the understanding of superpower retaliation to provocation, the assessment of formal models of crisis behaviour is a step beyond a debate over methods. It is interesting to note that, even before the specific performance levels of the respective models were considered, patterns emerged. This result is consistent with the argument that a form of rationality can be identified in the strategic interactions of the superpowers during international crises. In the debate over rational deterrence, some have argued, based on the record of case studies, that rational choice has virtually no explanatory power. If true, that conclusion would lead us to expect that the strategic choice in each case is likely to be unique and that no model can capture all cases, since misperception and other decision pathologies would produce extreme fluctuations in responses to provocations. In other words, it would be very *unlikely* that one model would predict the level of retaliation selected by the United States or the USSR, or the outcomes of serious disputes, more consistently than another.

Analysis of data on superpower interactions has suggested that OT and GRIT are more accurate reflections of crisis-based rationality. The performance of OT is quite impressive, especially given the concerns expressed by Beer (and others) about misperception. A complex model, OT *does* seem to have predictive power regarding retaliation by superpowers to provocation. These results can be related collectively to the outcomes that emerged from Axelrod's (1984) tournament for strategies designed to succeed in PD. He found that nice strategies – meaning those that started cooperatively – fared best in that setting. Strategies such as TFT therefore tended to become more common in the population over time. The results here suggest that strategies like GRIM, which is very punitive, are less likely to be implemented in superpower crises. Furthermore, the game-theoretic setting is one of

chicken, as opposed to PD, and a modified version of TFT outperforms the original. Perhaps OT is a more accurate reflection of bargaining under crisis conditions, whereas TFT tends to find favour in non-crisis settings.

Despite the performance of OT, TFT's record of success should not be downplayed. TFT has outperformed numerous rivals in a different setting, the iterated PD game. Thus it may be the case that strategies must be tailored to the type of game being played. To raise one possibility, it would be useful to study games in which at least one player has announced a strategy and the participants then interact over time. For example, it could be argued that Israel has been a GRIM player with the PLO. But Israel's interactions with Egypt and Jordan over border violations look more like those of GRIT. The choice of a strategy in an iterated game may depend on the nature of the adversary and a wide range of other factors.

EVALUATING CUMULATION

In accordance with the objective of developing integrative cumulation, the analysis described in this chapter will be compared to that of Goldstein and Freeman (1990) and Goldstein (1991), two frequently cited works on the postwar rivalry between the United States, the USSR, and China. They too attempted to assess the relative explanatory power of several game-theoretic models of strategic choice. As Goldstein and Freeman (1990, 1) note, "within each country there has been a running debate over the best strategy for eliciting desired responses from the other two countries ... Searching for these behavioral propensities in historical data ... is the key to strengthening the empirical footings of theories of cooperation." Like the present study, Goldstein and Freeman argue that traditional applications of game theory lack empirical content and that game theorists rely too heavily on "heuristic evaluations of arbitrarily selected real world events to demonstrate the validity of their arguments" (14). As a result, "we cannot say much about when, or if, the United States, the Soviet Union, or China has actually employed the kind of cooperative-reciprocal strategies these theories advocate." The purpose of the following section is to help improve our understanding of superpower rivalry and nuclear deterrence by emphasizing the extent to which the approach utilized in this chapter builds on Goldstein and Freeman's work in the area and provides a more accurate empirical assessment of game-theoretic models common to both projects.

Using a scale of net cooperation derived from a combination of data – the Conflict and Peace Data Bank (COPDAB), and the World Events

Interaction Survey (WEIS) – Goldstein and Freeman measured *high* and *low* levels of cooperation between the United States, the USSR, and China throughout the postwar period.³⁰ The objective was to catalogue *peaks* and *valleys* in mutual cooperation between each dyad in order to assess the relative explanatory power of four alternative strategies for transforming conflict into cooperation.

The four GRIT-type strategies are variations of Osgood's (1962) graduated reciprocation in tension reduction – TFT, GRIT, extended GRIT (EGRIT), and progressive GRIT (PGRIT). The differences are based on the number of conciliatory acts, the level (or intensity) of cooperation in each initiative, and the time frame (or pattern) within which the initiatives are sent. For example, "a country that employs TFT [as defined by Goldstein and Freeman] launches a single cooperative initiative at one point in time, while a country that employs PGRIT launches a continuous sequence of cooperative initiatives of escalating intensity" (Goldstein and Freeman 1990, 131). Plain GRIT entails a number of equally intense cooperative moves within a relatively short period of time, while EGRIT encompasses a strategy, similar to GRIT, whereby the cooperative moves are initiated over an extended period. Goldstein and Freeman evaluated the effectiveness of each strategy in terms of whether it produced "the greatest and most sustained increase in cooperative behaviour" following implementation of the strategy in question.³¹

Among the four alternatives, TFT was found to be the least effective, while strategies with repeated initiatives – GRIT, PGRIT and EGRIT – "have comparatively greater effects on superpower relations" (Goldstein and Freeman 1990, 134). GRIT and PGRIT were more successful than EGRIT at producing high levels of cooperation, although "an EGRIT strategy, if maintained indefinitely, did eventually drive U.S.-Soviet relations towards cooperation" as well. Goldstein and Freeman also found that, while GRIT produced the fastest increase, PGRIT produced "the greatest overall level of cooperation." In sum, the evidence suggested that "strategies using cooperative initiatives elicit like responses in the real world of great power politics ... When countries do something surprisingly cooperative in hopes of eliciting cooperation from another country, they generally meet with success. Hostile initiatives however ... tend to fail, eliciting hostile responses" (4). Goldstein and Freeman concluded that policy makers should use sustainable, cooperative initiatives in line with variations of GRIT and respond to cooperation in kind. Moreover, relatively high levels of cooperation should be preserved, even in the face of "substantial sustained hostility," in order to overcome forces (for example, bureaucratic inertia) that could stall implementation of the policy.

Although Goldstein and Freeman's (1990) study appears to provide important information about the behavioural dynamics of major powers involved in prolonged rivalry, several theoretical and methodological problems should be noted. To begin, Goldstein and Freeman claim to have evaluated the spectrum of strategies that could potentially transform conflict into cooperation. However, in reality they limited their investigation to variations on Osgood's GRIT-based alternative. They failed to assess the effectiveness of *coercive* threats – that is, GRIM-based strategies like deterrence, compellence, optimal threats, or any other variant of the “bullying” approach to crisis management. Consequently, the results were biased in favour of supporting one or another “nice” tactic.³² Although one nice strategy may prove to be more effective than another at promoting cooperation, in some cases, during an international crisis for example, they all may be less effective than coercive threats of retaliation.

The Cuban missile crisis serves as an excellent illustration. It was mutual fear of nuclear war, produced by the use of coercive threats on both sides, that ultimately pressured Kennedy and Khrushchev to resolve the immediate crisis, sign the Partial Test Ban Treaty, and improve the communications (hotline) link between Washington and Moscow. As Ray observes, “policymakers may find [Goldstein and Freeman's conclusions] too accepting of the assumption that cooperation in general is often a high-priority goal ... When the Chinese crack down on Tiananmen Square or the Soviets threaten the Baltics [or Iraq threatens Kuwait], is a surprising, unilateral cooperative move the best strategy?” (1991, 1494). In contrast to Goldstein and Freeman (1990), the present study assesses the *entire* spectrum of alternative strategies that may promote cooperation among rivals.

Although Goldstein and Freeman catalogued both peaks and valleys in cooperation between the United States, the USSR, and China since 1945, they focused exclusively on explaining the peaks and were concerned solely with assessing the ability of each strategy to promote long-term cooperation (1990, 42–4). Consequently, the other half of the puzzle remained unsolved: how to explain short-term cooperation between superpowers during the low points (that is, the valleys, or international crises). GRIT strategies aimed at fostering long-term cooperation may not be effective during a crisis, when decision makers are experiencing a threat to fundamental values, military hostilities are imminent, and there is a finite time available for a response. Under these conditions, conciliation is more likely to be perceived as a sign of weakness or capitulation. It is unlikely to convey the resolve required to protect vital security interests.

During noncrisis periods, on the other hand, cooperative initiatives generally are less costly ("talk is cheap"). The converse is also true; the use of coercive diplomacy during noncrisis periods is more likely to provoke hostilities and produce a sequence of mutual defections, not cooperation.

Clearly, both long-term (Goldstein and Freeman) and short-term (crisis) dimensions of the conflict-cooperation transition require attention for a comprehensive understanding of superpower rivalry. The present study complements that of Goldstein and Freeman by furnishing missing pieces of the cooperation puzzle. Unlike Goldstein and Freeman's study, the focus here is on determining the relative explanatory power of four alternative models of inducing cooperation (defined in terms of escalation control) between nuclear opponents at the brink of hostilities. Of course, given the potential impact of a mismanaged crisis on superpower relations, the findings generated in this study may even be more relevant to the issue of long-term cooperation than Goldstein and Freeman's (1990).

A third major problem is that the evidence Goldstein and Freeman offer to support the three GRIT-type strategies ignores an important dimension of U.S.-Soviet rivalry; specifically, GRIT, EGRIT and PGRIT probably were exercised *when there were sufficient assurances that each and every cooperative initiative would be met in kind*. It was extremely rare, given the Cold War mindset, for either side to initiate, let alone continue, cooperation without assurances that its acts would be reciprocated. Both Soviet and U.S. decision makers confronted numerous pressures to avoid even the appearance of unilateral concessions. The point is that these strategies cannot be identified for purposes of assessment solely on the basis of repeated cooperation, simply because such moves may have been a product of mutually assured cooperation, a more planned reciprocation by both sides that is not consistent with GRIT, EGRIT, or PGRIT.³³ It would have been appropriate for Goldstein and Freeman to provide evidence that the United States or the USSR intentionally and without knowledge of the opponent's likely response – indeed, regardless of that response – initiated a number of unilateral cooperative acts. However, they provide no indication that such actions were taken by either side.

Evidence presented by Goldstein and Freeman (1990, 135) to support GRIT and PGRIT also appears self-confirming, because the independent and dependent variables are indistinguishable. PGRIT, for example, is characterized as a strategy with a number of increasingly cooperative moves and is found to have a positive impact on the evolution of long-term cooperation among the superpowers. Perhaps

PGRIT (and other strategies with repeated initiatives) performed so well because the "progressively cooperative acts" (independent variable) that served as the defining characteristic of PGRIT encompassed the "sustained cooperation" (dependent variable) that the researchers intended to measure. It is not entirely clear how these variables can be separated in any meaningful sense without a precise measure of the level (or intensity) of cooperation in each initiative and of the leaders' perception of short and long time periods between initiatives. As Ray notes, most of the cases selected by Goldstein and Freeman "can be made to appear to conform to Osgood's general model rather easily with a judicious selection and weighting of events and time periods" (1991, 1494).

It is also interesting that Goldstein and Freeman emphasize the importance of their approach, in part, by criticising game-theoretic research and its "unitary actor" assumption. This criticism seems out of place for a study that depends so heavily on "events" data. After all, the information in WEIS and COPDAB is collected and organized on the basis of a unitary actor assumption, for example, levels of hostility and cooperation between state A, state B, state C, and so on. Furthermore, each of the strategies tested by Goldstein and Freeman is derived from rational choice premises. Regardless of the models being analyzed (TFT, GRIT, EGRIT, PGRIT, and the like), it is assumed that decision makers are capable of accurately perceiving the degree of cooperation (or noncooperation) in a series of initiatives, distinguishing between the available strategic options, and subsequently selecting a strategy that promises to elicit cooperation from the opponent. Although proponents of one strategy or another may disagree over which is most likely to promote short- or long-term cooperation, all sides assume that decision makers are logical and capable of living up to the demands of a rational process. If they were not, the entire research project would be moot. Once again, the present study provides a more reliable testing ground for alternative models of strategic choice, largely because decision making within the confines of a crisis is more likely to fall within the rational choice paradigm (Oneal 1988).³⁴

Several inferential problems with Goldstein and Freeman's use of daily events (as registered by COPDAB and WEIS) should also be noted. First, the United States, the USSR, and China are large countries with extensive organizations, large bureaucracies, and a myriad of standard operating procedures (SOPs), all of which serve to ease the foreign policy decision-making burden of leaders in noncrisis situations. Most activity initiated during these normal times, therefore, is likely to be a product of these bureaucratic and organizational forces. However,

Goldstein and Freeman fail to consider that “cooperation” may have resulted from these other forces. It is not clear how Goldstein and Freeman can evaluate the effectiveness of alternative strategies – all of which are based on rational choice assumptions – using data on events that may reflect forces that lie outside the rational choice (and to some extent realist) paradigm. Thus, inferences about the effectiveness of the strategies, based on the ebb and flow of cooperative events, are problematic. In contrast, the approach used in this chapter focuses on specific acts of provocation and retaliation when there is an explicit threat to values, a heightened probability of war, and time pressure, and when *the interest and attention of top level decision makers is clearly focused on the dispute in question*. Unlike research based on events data, the degree of cooperation and conflict during a crisis can be tied, more directly, to specific acts of provocation and retaliation and, therefore, to the strategies being analyzed.

Other complications associated with the use of events data that the present study avoids include the omission of certain kinds of important interactions among the players and the problem of adequately representing cooperation and conflict on a single scale. With respect to the first criticism, Goldstein and Freeman (1990, 40) concede that “some actions are secret; a country may be saying and doing all kinds of things in the public domain, but what the other countries care about are the tangible actions in the military and intelligence spheres, which remain hidden.” Of course, as Goldstein and Freeman argue, “the fact that public acts are an incomplete sample of the total (public and secret) acts biases the results *against* finding behavioral relationships ... [and] towards finding incorrect relationships only if the public acts consistently move in an opposite direction from the secret acts.” Since that is highly unlikely (41), any patterns identified in the relationship would provide evidence of the utility of events data.³⁵ This line of argument, however, is simply inadequate as a defence of COPDAB or WEIS. After all, the patterns that Goldstein and Freeman claim to identify – concerning either reciprocity or the explanatory power of different cooperation models – are more likely to be products of SOPS or bureaucratic politics, because most of the data represents activity occurring during normal (non-crisis) periods. On the other hand, the existence-of-patterns defence *is* appropriate for crisis-based research, because patterns of provocation and retaliation are less likely to be affected by other forces.

Goldstein and Freeman’s treatment of the Soviet invasion of Afghanistan, the Korean Airline incident, and the invasion of Grenada illustrates additional problems with using events data to gauge levels of cooperation and conflict. Based on the three data sets, the level of

hostility between the United States and the USSR during each of these cases was ranked *higher* than the Cuban missile crisis in 1962 (1990, 42–9). The notion that U.S.-Soviet relations during Afghanistan were more hostile (noncooperative) than in 1962 contradicts most historical accounts of the crisis; there was never any fear of nuclear war (or even conventional war) between the United States and the Soviet Union during the other cases. Clearly, events data is not sufficient to provide an accurate depiction of conflict and cooperation, unlike crisis-based data, which measures hostility in terms of provocation and retaliation (t_i and $q(t_i)$).

There is one final problem, related to the issue of reciprocity, that deserves mention. Proving that reciprocity between major powers takes place is essential for Goldstein and Freeman's investigation; if reciprocity did not occur, the issue of selecting and effectively utilizing specific GRIT strategies to illicit cooperation from an adversary would be irrelevant. Goldstein and Freeman observe that research on reciprocity, traditionally based on military expenditure and arms procurement data, generally has failed to support the reciprocity thesis. According to most studies on the subject, states are "largely driven by internal factors such as bureaucratic inertia [and] ignore changes in their adversaries spending" (1990, 25). In an effort to address this apparent anomaly, the authors note several problems with previous work on the subject and emphasize how earlier models of reciprocity were simply inconsistent with reality (for example, time series studies were aggregated over one-year periods and, therefore, hid a great deal of interaction; military expenditure data was often unreliable, especially for the USSR and China).³⁶ The objective was to develop an alternative approach towards testing reciprocity that supported the thesis. The issue here is not *whether* Goldstein and Freeman succeed or fail in their attempt to support the reciprocity thesis; it is that they *have to succeed* in order to justify their study. Given the overwhelming evidence against reciprocity noted by Zinnes (1980a, b), Bobrow (1982), and Patchen (1990), a more comprehensive defence is needed.

The purpose of the preceding review was to facilitate integrative cumulation on superpower rivalry and nuclear deterrence, in part, by emphasizing the extent to which the approach used in chapter 4 of this study provides a more accurate empirical assessment of game-theoretic models than Goldstein and Freeman (1990), a frequently cited work on the subject. Several propositions related to rational choice and coercive diplomacy more generally were evaluated using a crisis-based data

set. The objective was to identify patterns of response when one superpower was faced with a crisis resulting from a direct threat attributed to the opponent. The analysis of superpower rivalry moves beyond prior studies by reflecting a more basic concern for the interaction between aggregate research and the building of theory related nuclear deterrence and crisis management. Chapter 5 will include a case study of U.S.-Soviet interaction during the Yom Kippur War of 1973, one of the twenty-eight cases used in the superpower dataset. The purpose is to connect the aggregate results from chapters 4 and 5 with those obtained from an in-depth study of a single case, thus going beyond the information provided in the ICB case summaries.

5 U.S.-Soviet Rivalry and Nuclear Deterrence

As demonstrated in chapter 2, current scholarship on nuclear deterrence is primarily speculative and focuses on the behaviour *expected* from nuclear antagonists faced with the prospect of mutual annihilation. By comparison, insufficient attention has been paid to the relationship between expected behaviour and *actual* results, particularly in the context of superpower crises. This chapter explores two additional propositions central to the theory of nuclear deterrence. As in chapter 4, the objective is to determine whether states act according to the logic stipulated in the theory. The propositions to be tested will be identified in the first section, followed by operationalization of key variables and aggregate analysis in the second and third sections, respectively. A summary of the findings in the fourth section will set the stage for a more detailed case study of U.S.-Soviet interaction during the Yom Kippur War of 1973. The intention is to relate the aggregate results with those obtained from an in-depth study of a single crisis.

TESTING PROPOSITIONS

As Kugler (1984, 742) observes, “with impeccable logic, the rising spectre of nuclear devastation is linked to the preservation of peace. The very essence of this tightly structured argument is that, as the likelihood of virtual extermination increases, the probability that serious disputes will be resolved by nuclear war becomes exceedingly small. The terror created by the threat of nuclear devastation is,

therefore, the key to preventing the recurrence of major war." The neglected aspect, however, is whether the observed behaviour of states is consistent with that argument. Reproducible evidence on this linkage and others related to nuclear deterrence is not plentiful.

Observers agree, of course, that the deterrent role played by nuclear weapons cannot be tested directly in the absence of nuclear war. It is possible, however, to assess the role of these forces indirectly by focusing on some of the attendant propositions from deterrence theory (Bueno de Mesquita and Riker 1982; Kugler 1984). For example, if the logic of nuclear deterrence is sound, *the violence experienced by the participants in a dispute should decrease as the threat posed by nuclear weapons increases* (proposition 1, or P1). The more acute the threat of nuclear use, in other words, the more likely that decision makers will search for a solution that promotes de-escalation of hostilities, in order to avoid the worse payoff for both: mutually assured destruction (MAD).

Some investigations, reviewed in chapter 3, have reported positive findings for P1, suggesting the need for replication. Bueno de Mesquita and Riker (1982) found that nuclear weapons appear to constrain the intensity of conflict, and Weede (1981, 1983) produced results favourable to extended deterrence. These studies currently stand in isolation. By contrast, Huth and Russett (1984, 1988, 1990) confirmed Russett's (1963) initial findings concerning economic (as opposed to nuclear) linkages and successful deterrence. Other propositions tested by Huth and Russett (1988) have not, however, been dealt with elsewhere.

Among the negative findings, the most consistently reported is that possession of nuclear weapons does not seem to facilitate deterrence. Organski and Kugler (1980), Kugler (1984), Huth and Russett (1984, 1988, 1990) and Betts (1988) have all obtained that result through different approaches. This negated linkage leads naturally into the questions raised earlier about validity and reliability.

Despite the centrality of P1, however, deterrence encompasses more than simply avoiding war; in fact, the treatment of the dependent variable has been biased. Most studies have equated deterrence failure with the outbreak of war and other forms of violence. However, in assessing outcomes, a greater role should be reserved for the achievement or denial of the defender's and attacker's policy interests (Kugler 1984; Huth 1990; Huth and Russett 1990). Clearly, if states manage to elude war by threatening the use of nuclear weapons, but fail to obtain policy objectives in that process or to prevent the adversary/attacker from obtaining its goals, nuclear deterrence cannot be regarded as completely successful. As Huth and Russett (1990) argued, cases of successful deterrence should include instances in

which the attacker did not resort to military force and *did not coerce the defender into capitulating to its demands*. Similarly, deterrence failures should include cases in which *the attacker either attained its policy goals* (under the threat of force) or resorted to the sustained use of force. It is not only the costs of violence that are considered in a decision maker's utility calculus, but the costs associated with overall victory and defeat. Of course, a nuclear state might fulfil its goals during a dispute, but at the cost of escalating a crisis to the level of war, and that also would be unsatisfactory.

The point is that although the occurrence or nonoccurrence of war and major hostilities is an important indicator, valid testing requires a research design that specifies a range of dependent variables, including goal achievement, in order to determine whether interactions among states confirm the logic derived from the theory (Harvey 1995). Specifically, *nuclear states generally should prevent nonnuclear aggressors from achieving their policy interests in crises and achieve draws when confronted with other nuclear powers* (proposition 2, or P2). Together, P1 and P2 offer the most straightforward approach to assessing theoretical predictions.

OPERATIONALIZATION OF VARIABLES

The basis of evidence, as in chapter 4, consists of twenty-eight international crises between 1948 and 1988 in which one or both superpowers were involved as crisis actors. It was expected that as the threat posed by nuclear weapons increases (the independent variable), the violence experienced by the participants in the dispute would decrease (P1) and the likelihood that policy objectives of the nuclear nation are satisfied would increase (P2).

Independent Variables

The independent variable (nuclear threat) was operationalized using a variant of the severity index developed by Brecher and James (1986).¹ The new index, which will be referred to as the severity of threat (ST) index, does not measure the overall severity of a crisis, but instead assesses the severity of the threat *experienced by the nuclear power*. The logic of nuclear deterrence stipulates that nuclear retaliation becomes more likely, and therefore more credible, as a threat to a nuclear state becomes more acute, presumably because a severe threat is required before the decision makers contemplate the use of nuclear weapons. For example, a threat of nuclear retaliation would tend to be more credible during a crisis in which a state's existence is threatened. On

the other hand, a similar threat of nuclear retaliation may not be as credible in response to an attack on an ally or client state (for example, an American threat of nuclear retaliation in response to a Soviet menace to Western Europe). The objective is to measure the relationship between the severity of the threat that, according to the theory, is expected to restrain conflict behaviour, most notably the level of violence experienced by the actors.

Six indicators were used to measure the overall severity of a crisis: the extent of superpower involvement, the level of violence, the number of crisis actors, geostrategic importance, heterogeneity, and issue area. Three modifications were required for present purposes. First, since superpower involvement applies (by definition) to all twenty-eight crises, it was not used to establish the severity of a threat. Second, "violence" represents the dependent variable in this study and, therefore, could not be included in the ST index. Finally, very few of the twenty-eight crises listed in Table 3.1 involved more than three crisis actors – the actors were usually a superpower initiator, target, and superpower defender. Consequently, variations in the number of actors would have had no appreciable impact on ST and, therefore, were excluded from the analysis. Thus three indicators were preserved: (1) geostrategic importance, (2) heterogeneity, and (3) the issues involved. Three other indicators were also included: (4) the gravity of a threat to superpower values, (5) the perceived threat to superpower influence, and (6) the crisis trigger, or provocation. Each will be described in turn.

Geostrategic importance, which may vary over time, refers to the location of a crisis "in terms of its natural resources, distance from major power centers ... [and] the significance of the location and resources for one or more international systems" (Brecher and James 1986, 36). For example, "[o]il and uranium-producing regions acquired greater salience since the 1950s; coal-producing regions became less significant. Key waterways and choke points such as the Suez Canal have retained their (geostrategic) relevance over the decades. A combination of assets enhances geostrategic salience. For instance, the strait of Hormuz provides access and egress from a region that, in the 1970s, produced 40% of the world's oil supply, with vast reserves to increase its geostrategic salience." A five-point ordinal scale (Table 5.1) represents the different degrees of geostrategic importance: relevance (1) to one subsystem, (2) to more than one subsystem, (3) to the dominant system and one subsystem, (4) to the dominant system and more than one subsystem (for example, the Berlin crises of 1948 and 1961), and (5) to the global system (for example, the Cuban missile crisis of 1962, the Suez crisis of 1956, the

Table 5.1
Degrees of Geostrategic Importance

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
5 Global system	1.000	1.000
4 Dominant system and more than one subsystem	0.750	0.682
3 Dominant system and one subsystem	0.500	0.414
2 More than one subsystem	0.250	0.189
1 One subsystem	0.000	0.000

Note: An explanation of the linear and exponential threat ratings, the assumptions behind them, and the functions used to generate linear values and their exponential equivalents is given at the end of this section on independent variables.

Table 5.2
Degrees of Heterogeneity

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
5 Military, economic, political, and cultural differences among adversaries	1.000	1.000
4 Three attribute differences	0.750	0.682
3 Two differences	0.500	0.414
2 One difference	0.250	0.189
1 None	0.000	0.000

Six Day War of 1967, and the Yom Kippur War of 1973). A crisis whose outcome is relevant to the structure of the global system, that is, to the global balance of power, poses a much greater threat to a superpower than one in which the outcome effects only one subsystem.

Heterogeneity refers to the number of "attribute differences among adversaries" in a crisis. The attributes in question are military capabilities and economic, political, and cultural differences among adversaries (Table 5.2). Differences in all four attributes imply a high level of heterogeneity, while differences in only one or in none represent a low level of heterogeneity. Brecher and James (1986, 38) note that "more attribute differences among adversaries in an international crisis indicate more cleavage and, therefore, greater severity." Several case studies are cited as evidence that multiple cleavages make political accommodation more difficult and military solutions more likely (137).²

Table 5.3
Issues

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
5 Three or more issues	1.000	1.000
4 Two issues including at least one military-security issue	0.750	0.682
3 One military-security issue alone	0.500	0.414
2 Two issues other than military-security issues	0.250	0.189
1 One nonmilitary-security issue	0.000	0.000

Issues are divided into four general categories: military-security issues (for example, the incorporation of territories, a dispute over borders, changes in the military balance, interstate war); political-diplomatic issues (for example, issues related to sovereignty, hegemony, and international status and influence); economic-developmental issues (for example, nationalization of property, control over raw materials, economic pressures such as boycotts and sanctions, foreign exchange problems); and cultural-status issues (for example, ideological or religious disputes, challenges to nonmaterial values). Brecher and James assume that "a crisis over a military-security issue alone indicates more severity than a crisis concerned with any other issue-area" (1986, 137). Based on this assumption, a five-point scale of intensity was developed representing (1) one nonmilitary issue (the lowest level of severity); (2) two issues other than military-security issues; (3) a military-security issue alone; (4) two issues, including at least one military-security issue; and (5) three or more issues, including at least one military-security issue (the highest level of severity) (Table 5.3).

The gravity of a threat to superpower values, the fourth indicator, refers to the value that the crisis actor perceives to be the object of gravest threat during a crisis. Threats range from the most severe – a threat to a country's existence or to the survival of a population, genocide, total annexation, or occupation (for example, Israel and the Six Day War of 1967) – to the least severe (for example, a threat of economic retaliation, a limited threat to a country's property, or blocked access to resources or markets). To measure severity, the following nine-point scale was used: (1) no threat, (2) a limited threat to population and property, (3) a threat to a social system, (4) a threat to economic interests, (5) a threat to a political system, (6) a threat to territorial integrity, (7) a threat to influence in the international system, (8) a threat of grave damage, and (9) a threat to the existence of the state (Table 5.4).

Table 5.4
Gravity of Threats to Superpower Values

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
9 Threats to existence	1.000	1.000
8 Threats of grave damage	0.875	0.835
7 Threats to influence	0.705	0.682
6 Threats to territory	0.625	0.523
5 Threats to a political system	0.500	0.414
4 Threats to an economic system	0.375	0.297
3 Threats to a social system	0.250	0.189
2 Limited threats	0.125	0.091
1 No threats	0.000	0.000

Indicator 5, perceived threats to superpower influence, is essentially an extension of the third scale-point for indicator 4, the gravity of a threat to superpower values. If a superpower perceives a threat to its influence in the international system – a common perception on the part of both superpowers in most crises throughout the Cold War – it is important to identify the type of influence that was threatened. For example, a threat to American influence within the Western bloc would pose a much greater problem for American decision makers than a threat to American influence in Eastern Europe. The following scale represents the ordinal severity points for the influence indicator: (1) no threat to influence of a superpower (low severity), (2) a decline in superpower influence within the adversary's bloc, (3) a decline in superpower influence over nonaligned groups, (4) a decline in superpower influence over nonbloc client states, (5) a decline in superpower influence within its own bloc, and finally, (6) an adverse change in the global balance of power (high severity) (Table 5.5).

The final indicator is the trigger of the crisis, defined as the "specific act(s), event(s), or situational change(s) which generated, for the decision maker(s) of the crisis actor, a perception of threat to basic values, time pressure, and the likelihood of involvement in military hostilities" (Wilkenfeld, Brecher, et al. 1988, 2). The more violent the trigger, the more severe the provocation or threat. A crisis triggered by an indirect violent military act, as in the Soviet crisis triggered by Israel's invasion of the Golan Heights in 1967, for instance, would pose a greater threat to the Soviet Union than a crisis triggered by a mere verbal act, for example, by the War of Attrition in 1970. The 15 scale points for the trigger range from the most severe threat (a nuclear strike) to the least severe (cooperation) (Table 5.6).

Table 5.5
Perceived Threats to Superpower Influence

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
6 Adverse change in global balance of power	1.000	1.000
5 Decline in superpower influence within its own bloc	0.800	0.741
4 Decline in superpower influence over nonbloc client states	0.600	0.516
3 Decline in superpower influence over nonaligned groups	0.400	0.319
2 Decline in superpower influence within the adversary's bloc	0.200	0.149
1 No threat to influence of superpower	0.000	0.000

Table 5.6
Crisis Triggers

<i>Criterion (s)</i>	<i>Threat Rating (exponential scores)</i>
15 Nuclear strike	1.000
14 Full-scale conventional war	0.903
13 Indirect full-scale conventional war	0.811
12 Serious clashes	0.724
11 Indirect serious clashes	0.641
10 Minor clashes	0.561
9 Indirect minor clashes	0.486
8 Multiple triggers, including violent military acts	0.414
7 Nonviolent military act	0.346
6 Multiple, including nonviolent military acts	0.281
5 Economic act	0.219
4 Political act	0.160
3 Verbal act	0.104
2 No action	0.051
1 Cooperation	0.000

Note: Only exponential scores were included in this table because they represent the stronger (more realistic) assumption about threat levels.

The scale points for each of the six indicators of ST were assigned a numerical rating ranging from 0.00 for a very low threat to 1.00 for a very high threat. The indicators and values assigned to their scale points are displayed in Tables 5.1 through 5.6. Using these parameters, the overall ST for any given crisis ranged from a high of 6.00 (a score of 1.00 on all six indicators) to a low of 0.00. Although a more

elaborate weighting scheme could be devised to establish the relative potencies of each indicator, this is not essential for the present study. The objective here is to offer preliminary support for the two propositions, P1 and P2.

Both linear and exponential threat ratings were used to test two alternative assumptions regarding the difference in severity between consecutive scale points for each indicator. The linear rating assumes that as we move from 0.00 to 1.00 on the scale, there is a uniform increase in severity. For instance, the difference between scale points 2 and 1 on the geostrategic importance indicator – that is, more than one subsystem and one subsystem, respectively – is equal to the difference between scale points 4 and 3; there is a uniform variation of .250 between any two consecutive linear points. The function used to generate the linear values is as follows:

$$(1) L = (s - 1) / (n - 1),$$

where L = the linear scale point; s = the scale point number; and n = the total number of scale points. On the other hand, an exponential rating assumes that there is a larger difference in severity between consecutive scale points as we move closer to a score of 1.00. For example, the difference in severity between scale point 5 (the global system, Table 5.1) and scale point 4 (the dominant system and more than one subsystem) is greater than the difference between points at the lower end of the scale. The rationale for an exponential function is that “more cooperative behaviour leads to diminishing returns at the interstate level” (James 1988, 20). Threat, in other words, tends to increase at a faster rate as we move up each scale. The function used to transform the intervals between the scale points into their exponential equivalents is as follows:

$$(2) E = e^{0.693(L)} - 1,$$

where E = the exponential scale point; and L = the linear scale point.

Dependent Variables

Violence, the first dependent variable, was measured with four indicators: (1) the major response by the superpower/defender to the trigger of the crisis, (2) the principal techniques of crisis management used by the superpower/defender, (3) the intensity of violence in the crisis, and (4) the centrality of violence in the crisis.³ Like ST, the scale points for each indicator were assigned both linear and exponential

Table 5.7
Major Responses

<i>Criterion (s)</i>	<i>Threat Rating (exponential scores)</i>
15 Nuclear strike	1.000
14 Full-scale conventional war	0.903
13 Indirect full-scale conventional war	0.811
12 Serious clashes	0.724
11 Indirect serious clashes	0.641
10 Minor clashes	0.561
9 Indirect minor clashes	0.486
8 Multiple responses, including violent military acts	0.414
7 Nonviolent military act	0.346
6 Multiple, including nonviolent military acts	0.281
5 Economic act	0.219
4 Political act	0.160
3 Verbal act	0.104
2 No action	0.051
1 Cooperation	0.000

Note: Only exponential scores were included in this table because they represent the stronger (more realistic) assumption about threat levels.

Table 5.8
Crisis-Management Techniques

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
4 Violence	1.000	1.000
3 Multiple techniques involving violence	0.666	0.587
2 Nonviolent military techniques	0.333	0.259
1 Negotiation, arbitration, mediation, nonmilitary pressure	0.000	0.000

ratings that ranged from 0.00 (very low level of violence) to 1.00 (very high level). The four indicators, and the values assigned to their scale points, are displayed in Tables 5.7 through 5.10, with overall levels of violence ranging from a high of 4.0 (a score of one on each indicator) to a low of 0.0.

The second dependent variable is the achievement (or nonachievement) of the defender's goals or objectives in the crisis. Both subjective and objective measures are used. The defender's level of satisfaction, upon termination of the crisis, represents a subjective assessment of the outcome. The following scale points are used: (1) defender dissatisfied,

Table 5.9
Intensity of Violence

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
4 Full-scale war	1.000	1.000
3 Serious clashes	0.666	0.587
2 Minor clashes	0.333	0.259
1 No violence	0.000	0.000

Table 5.10
Centrality of Violence

<i>Criterion (s)</i>	<i>Threat Rating</i>	
	<i>Linear (L)</i>	<i>Exponential (E)</i>
4 Violence exclusive	1.000	1.000
3 Violence central	0.666	0.587
2 Violence minor	0.333	0.259
1 No violence	0.000	0.000

initiator satisfied (least preferred); (2) defender dissatisfied, initiator dissatisfied; (3) defender satisfied, initiator satisfied; and (4) defender satisfied, initiator dissatisfied (most preferred). An objective assessment of crisis outcomes is measured in terms of (1) defeat, (2) stalemate or compromise, and (3) victory.

ANALYSIS OF DATA

Proposition P1 is supported by the data in Table 5.11. Of the thirteen crises in which the superpower defender experienced a very high threat (4.1 to 6.0), in ten (77 percent) the violence among actors in the dispute, including clients of each superpower, was minimal (0.0 to 0.88) ($\tau b = -.39$; $p < 0.09$; $N = 28$).⁴ Furthermore, of these thirteen cases, the participants experienced high levels of violence in only two (15 percent). Of the twenty cases that generated low violence scores, in ten (50 percent) the defender experienced a high threat, with the remaining ten cases exhibiting threat levels between 2.1 and 4.0. Similar results were obtained using the exponential scale points.⁵

Other measures of association appropriate for ordinal variables provide additional support for P1. As expected, there is a negative association between the severity of the threat and the violence experienced

Table 5.11
Provocation and Retaliation in Foreign Policy Crises, 1948-88

Case	Provocation by	Retaliation by	Threat Level		Response Level	
			Linear	Exponential	Linear	Exponential
1	USSR	USA	4.43	4.07	0.69	0.54
2	USA	USSR	4.21	3.98	0.76	0.60
3	USA	PRC/USSR	4.30	3.93	3.36	3.28
4	USSR/PRC	USA	3.13	2.76	0.76	0.60
5	USA	USSR	4.61	4.32	3.43	3.35
6	USA	PRC/USSR	4.93	4.71	2.21	1.92
7	USA/Taiwan	PRC/USSR	3.81	3.45	1.05	0.90
8	USA/Taiwan	PRC/USSR	3.81	3.45	1.05	0.90
9	USSR	USA/Britain/ France	4.53	4.23	0.69	0.54
10	USA/Turkey	USSR/Syria	2.41	2.12	0.48	0.36
11	USA/Taiwan	PRC/USSR	4.03	3.64	0.14	0.10
12	USSR	USA/Britain/ France	3.59	3.36	0.55	0.42
13	USA/Britain/ France	USSR	4.68	4.39	0.36	0.28
14	USA	USSR/Cuba	3.41	3.15	0.14	0.10
15	USSR/GDR	USA/FRG	4.91	4.70	0.76	0.60
16	USSR/Cuba	USA	4.68	4.44	0.76	0.60
17	USA	USSR/Cuba	3.80	3.59	0.69	0.54
18	USA/Belgium	USSR	4.19	3.82	1.14	0.88
19	USA	USSR/North Vietnam	3.49	3.24	3.02	2.87
20	USSR	USA/Israel	4.86	4.64	0.76	0.60
21	USSR/Egypt	USA/Israel	3.78	3.32	1.69	1.39
22	USSR	USA	3.30	3.01	0.14	0.10
23	USSR/Egypt/ Syria	USA/Israel	4.86	4.64	0.69	0.54
24	USA	USSR/ Ethiopia	2.05	1.70	2.21	1.92
25	USA	USSR/ Grenada	2.94	2.72	2.88	2.75
26	USSR/Grenada	USA	1.93	1.76	1.79	1.57
27	USA	USSR/ Nicaragua	3.61	3.38	0.48	0.36
28	USSR/ Nicaragua	USA	2.43	2.17	0.36	0.28

Note: Crises are identified in Table 3.1.

by the participants in the crisis ($\gamma = -0.58$): an increase in severity often produced *lower* levels of hostility. Many of these findings are consistent with those generated from the game-theoretic analysis of provocation and retaliation developed by James and Harvey (1992). They found that as the intensity of a provocation increased, the relative severity of the retaliation decreased ($r = -0.7633$, $p < 0.001$, $N = 28$). Intense provocations, therefore, are unlikely to generate an equal or more intense response from the other superpower, presumably because of the mutual fear of escalation. Similarly, Brecher (1993) found a link between military power and crisis behaviour by states that possess nuclear capability. Those with large stockpiles are likely to be very prudent in using violence to cope with crises, because of the danger of rapid and destructive escalation. These patterns of behaviour among nuclear powers faced with intense crises support predictions of deterrence theory.

Also, violence tended to be higher in situations in which the threat was neither very high nor very low, which is consistent with conventional wisdom on the effects of grey areas in deterrence situations. Crises in areas of the world that are not geostrategically salient, for example, but nevertheless have a great deal of importance with respect to superpower influence, both within the region and globally (for example, Vietnam and Korea), would tend to result in substantially higher levels of violent behaviour.

Although P2 was supported by the data, the strength of the relationship depended on whether objective or subjective assessments of the outcome were used. When the outcome was measured in terms of defeat/compromise/victory, the superpower defender achieved outright victory in only six (46 percent) of all high threat cases. The remaining 54 percent ended either in a stalemate/compromise (24.7 percent) or a defeat (29.3 percent). A very high threat was present in six of the seven cases in which the defender was defeated (85.7 percent). When the outcome was measured in terms of the defender's satisfaction level, the results were somewhat different. Of the thirteen cases in which the superpower defender experienced a very high threat, in nine (69 percent) a very high satisfaction level was achieved. Moreover, in 58 percent of the cases where the superpower achieved the best outcome (that is, the defender was satisfied and the attacker was dissatisfied), a very high threat was present. Once again, linear and exponential measures produced similar results.⁶

Among the more interesting findings is that deterrence success depends on the identity of the defender. The results demonstrate that the United States (as defender) was more successful than the Soviet Union in obtaining its objectives in high-threat situations. Of the cases

in which the United States experienced a high threat, its goals were achieved in 80 percent. On the other hand, the Soviet Union achieved its goals in only 43 percent of its high-threat cases. Similar results were obtained using the exponentially derived *sr* index (77 percent for the United States versus 60 percent for the Soviet Union). This pattern is consistent with the findings of Wilkenfeld and Brecher: "goal achievement was uniformly higher for the United States (72% of its crises), regardless of the type of issue area involved. This tendency was accentuated in the subset of U.S. cases involving confrontation with the USSR – 81%" (1979, 20).

The findings generally supported both proposition P1 and P2. As the threat associated with potential nuclear retaliation increased, the violence experienced by participants tended to decrease. The severity of the nuclear threat also affected the likelihood of goal achievement, although this result depended on whether an objective or subjective assessment was used. In terms of satisfaction, the evidence supported P2, while the results were much weaker when goal achievement was assessed in terms of defeat and victory. In addition, the assertion that a stalemate or compromise would result from situations that involved mutual deterrent threats was not supported by the evidence. Winners and losers still emerged in these cases, and the United States tended to be more successful than the Soviet Union at achieving objectives in high threat situations.

SUMMARY AND CASE STUDY

The debate over rational deterrence – especially at the nuclear level – cannot move forward without an authoritative body of evidence.⁷ Although the focus of the preceding investigation has been on testing deterrence in the nuclear realm, the same propositions apply to crises involving nonnuclear states faced with the prospects of large-scale conventional war. Despite efforts to identify a suitable list of cases, however, the dominant testing strategy appears to have come up short, largely due to the inability of scholars to be definitive about the myriad of coding decisions required for each case. In the conviction that there are alternative approaches, twenty-eight cases of superpower rivalry were used to evaluate two propositions, P1 and P2, derived from nuclear deterrence theory. Predictions of rationality that underlie nuclear deterrence theory appeared to be supported by superpower behavioural patterns during international crises since 194⁸. Although the objective was to provide a more clearly identified and common purpose to facilitate convergence of the research program, the results presented in this and the preceding chapter cannot be definitive

without a detailed assessment of individual cases. With this in mind, the next section explores U.S.-Soviet interaction in the Yom Kippur War (1973), as an approach towards linking aggregate data on nuclear deterrence with data obtained from an in-depth study of decision making in a single crisis.

Yom Kippur, 1973

The Yom Kippur War was a culmination of increasing Egyptian frustration over Israeli occupation of Egyptian territory since the Six Day War of 1967. The status quo was unacceptable to Egypt, to say the least, and on 6 October 1973, President Sadat spearheaded a joint Egyptian-Syrian attack against Israeli troops in the Sinai and Golan Heights. The main participants in the crisis were Israel, Egypt, Syria, the United States, and the Soviet Union, but given the specific concerns of the present inquiry only U.S. and Soviet behaviour will be investigated here. The analysis is not intended, in any way, to be a comprehensive review of the crisis, but only a preliminary report of U.S.-Soviet behaviour with specific emphasis on the extent to which their actions are consistent with the coding decisions of chapters 4 and 5 and with theoretical predictions.

At the superpower level, the status quo was *détente*, and there were ongoing efforts to maintain it. Both the United States and the Soviet Union sought to preserve *détente* by downplaying potential areas of conflict, especially in the Middle East. In fact, in the final communiqué of the U.S.-Soviet summit issued on 24 June 1973, only 87 out of 3200 words dealt with the region (Freedman 1975, 113). American officials, convinced that Israel's military superiority and deterrent threat would be sufficient to prevent a crisis from escalating out of control, were basically content with the status quo. The reference point for American leaders, therefore, included a strong relationship with Israel, an improving American-Arab relationship, and a declining level of Soviet influence in the Middle East. As long as Israel's position of strength vis-à-vis other states in the region was preserved, a low-key approach to the crisis was preferred. President Nixon's objective was to maintain the American sphere of influence, preserve *détente*, and "achieve peace without losing the support we had been able to build up in both the Arab and Israeli camps" (Nixon 1978, 924). Of course, any deterioration in American relations with Israel would be viewed as a serious loss, requiring a more direct approach.

To accomplish these objectives, the initial American reaction to the outbreak of hostilities was cautious and entailed a relatively low level of risk: "we had to protect our country's ability to play its indispensable

role as the guarantor of peace ... We would need to show that we were determined to prevent Soviet intervention, but we had to do so in a low-key way, conveying confidence without weakness" (Kissinger 1982). Nixon's official response was free of promises of increased military support for Israel; it was even free of blame. However, two key elements conditioned American actions later on in the crisis: the threat of Soviet intervention and Israel's changing military position in the war. For Nixon and Kissinger, Soviet intervention or a worsening Israeli position would result in significant losses for the United States and would require a more direct response.

By 12 October, it became clear to all parties that a swift Israeli victory was unlikely. Although Israeli officials had advance knowledge of the initial attack, they decided not to be the first to move, gaining the political benefits that went along with a defensive position. The price they paid for that kind of restraint, however, was quite heavy, with positions being abandoned in the Golan Heights and Sinai and deployments, by Egypt, of an additional one hundred thousand troops around the Suez Canal (Stoessinger 1993). There were heavy casualties on both sides in some of the largest tank battles since World War II. Soviet ground-to-air missiles were especially lethal against the Israeli air force, the centrepiece of Israel's military strength. As casualties mounted, the United States and the Soviet Union became increasingly involved with supplying ammunition, tanks, and planes to each side. The Americans also obtained evidence that the Soviets were pressuring Jordan to enter the battle and were stepping up their airlifts of arms and supplies to Syria (Kissinger 1982, 497).

In response to these events, Kissinger and Nixon modified their strategy. Previous calls for a cease-fire based on the status quo were replaced by demands for a "cease-fire in place," while shipments of arms to Israel were increased (Quandt 1977, 176). Both Nixon and Kissinger began to treat the high probability of an Israeli defeat as a certainty. Fearing substantial losses, the American position shifted from a minimal level of involvement to a conspicuous stance in support of Israel, despite the implications for U.S.-Soviet relations and the possibility of more direct superpower military confrontation.

During the second week of the war, Israel began to gain some ground, surrounding Egypt's Third Army in Sinai and cutting off supplies of food and water. In response to this turn of events, Sadat formally requested the help of both superpowers to enforce the cease-fire and prevent Israel from destroying the Egyptian army. After a definitive American refusal, a letter from Brezhnev was delivered to the White House stating that if the cease-fire was not observed by Israel and the United States refused to cooperate in a joint task force, the

Soviet Union "would take appropriate steps unilaterally" (Golan 1977, 121). The American reaction was to immediately move to DEFCON 3 – a nuclear alert involving partial mobilization of American forces. This decision involved the highest amount of risk. The probability of military escalation, and even a nuclear exchange, increased as both Soviet and American forces heightened their alert status. The reason for the "costly signal" was Kissinger's strong belief that unilateral intervention by the Soviets would give them greater control over the outcome of events in the region and would make officials in Moscow less willing to compromise in light of their new position. U.S. officials responded to what they perceived as a clear, and potentially irreversible, loss – particularly in light of the recent military gains by Israel.

The Soviets were at least as sensitive as the United States to any negative deviations from the status quo. Soviet interests and credibility in the region, after all, would have been seriously undermined if Egypt and Syria were to suffer, once again, at the hands of Israel and the United States. At the same time, serious efforts were being made to prevent a worsening of relations with the United States. As a consequence, the Soviets were forced to strike a difficult balance between preserving gains from détente and remaining committed to sending substantial military and economic aid to Egypt and Syria.

In line with this two-track strategy, the initial Soviet response was to condemn Israel and press diplomatically for a return to pre-1967 borders. Soviet media coverage of the war was restrained, ships were withdrawn from the area to emphasize the regional nature of the crisis, and diplomats were immediately dispatched to Egypt to convince Sadat of the benefits of a cease-fire and of using less direct means of confrontation (for example, the "oil weapon"; Golan 1977, 74). In addition, officials in Moscow informed Sadat that he could expect only indirect Soviet assistance in case of war, in order to avoid a direct U.S.-Soviet clash (14). Even when hostilities broke out, the Soviets sought to limit escalation and preserve both détente and their position with their Arab allies.

As was the case with the United States, the Soviet strategy became more risk acceptant as the fighting escalated and as Egypt began to lose ground. Shortly after the first signs of a Syrian retreat (8 October) and partially in response to TASS reports of increased American intervention, the Soviets stepped up their own airlifts of supplies. When Sadat requested joint superpower intervention to prevent the destruction of the Third Army, the crisis entered its final stage with the Soviets facing an almost certain loss. In response, Brezhnev threatened unilateral action in the form of Soviet troop deployments to stop the fighting. Clearly, the Soviet threat was intended, in part, to force the

United States to pressure Israel into accepting the cease-fire, even though the risks of escalation were high. But, in line with nuclear deterrence theory, the Soviets were not willing to risk a potentially greater loss in the form of a military confrontation with their nuclear rival, even though they faced an almost certain loss of influence and credibility in the region.

Both superpowers demonstrated an aversion to losses by adopting risk-acceptant behaviour. The American decision to move to DEFCON 3 demonstrated the desire to react forcefully and decisively to deter Soviet intervention. Similarly, Soviet involvement in the war at increasingly higher levels and culminating in an explicit threat to intervene unilaterally confirmed that Soviet leaders were willing to risk escalation to protect their influence in the Middle East, but only up to a point.

6 Nuclear Stability after the Cold War

The behavioural patterns of the United States and the Soviet Union during international crises from 1948 to 1988 are mostly consistent with expectations of rationality derived from nuclear deterrence theory. The results have emerged from two alternative testing strategies, presented in chapters 4 and 5, respectively.¹ Several questions about nuclear stability in a post-Cold War world, noted at the outset of this project, are addressed in the first section of this chapter: specifically, Will the collapse of the Soviet Union create a more stable or a more hostile nuclear environment as we approach the twenty-first century? Will the shift away from power bipolarity intensify or diminish the proliferation of nuclear weapons in the Middle East and other parts of the developing world? If proliferation intensifies, will the crisis management behaviour of prospective nuclear states mirror the patterns exhibited by the United States and the USSR throughout the Cold War? The second section evaluates, in more general terms, the ongoing importance of research on U.S.-Soviet crises and the continuing relevance of realist theories, like deterrence, to the study of international politics and foreign policy. The conclusion summarizes the contributions of the project and returns to the issue of integrative cumulation in IR.

PLAYING BY RULES: BOUNDED RATIONALITY IN NUCLEAR CRISES

Gaddis' (1986, 99) observations are appropriate as a starting point: "the post-World War II system of international relations, which nobody

designed or even thought could last for very long, which was based not upon the dictates of morality and justice but rather upon an arbitrary and strikingly artificial division of the world into spheres of influence, and which incorporated within it some of the most bitter and persistent antagonisms short of war in modern history ... survived twice as long as the far more carefully designed World War I settlement."² Every U.S.-Soviet confrontation was terminated without a shot being fired, though few IR scholars in the late '40s and '50s were confident that a thermonuclear holocaust could be avoided; this lack of confidence was widespread among most students of international affairs throughout the post-World War II era. "It is most unlikely," Kahn argued in 1961, "that the world can live with an uncontrolled arms race lasting several decades ... We are not going to reach the year 2000 - and maybe not even 1965 - without a cataclysm" (574). These widespread expectations of disaster were reinforced by scholarship that continued to paint a picture of the incompetent, pathological decision maker with a penchant for misperceiving or miscalculating the intentions of the adversary due to some innate inability to cope with stress and anxiety during crises. Contrary to the image of a pathological leader routinely depicted by critics of rational choice, a radically different image of crisis decision making emerged from empirical observations noted in this study.

While the content of provocation and retaliation may have changed from one dispute to the next, strategic choice was more predictable than might have been imagined.³ Despite the stress experienced by decision makers in each case, the powerful political and psychological pressures to prevail every time, mutual distrust, ideological differences, and many other obstacles to cooperation in security matters, both sides developed strategies to manage their rivalry without crossing the brink. It appears, therefore, that crises involving nuclear states are more likely to exhibit characteristics of restraint, as Brodie predicted (1959a,b), because each side understands the dangers of any action that could raise tensions to the level of war. Leaders have grasped the general requirements for crisis stability and have learned several vital "rules of prudence" (Craig and George 1990).⁴ It was not just luck that prevented hostilities from breaking out between the United States and the USSR but the opportunities and constraints embedded within superpower rivalry, bipolarity, nuclear deterrence, and the logic of mutual assured destruction.

The fact that the record of superpower activity generally supports the presence of certain strategies, and not a random distribution of retaliation to provocation as implied by political psychologists, lends some credibility to the rational choice perspective. U.S. and Soviet officials did not appear to be overrun with doubts and uncertainties,

excessively concerned about cognitive consistency, reluctant to make irrevocable policy choices, or constantly engaged in any of the defensive avoidance techniques noted in the literature. On the contrary, both sides responded to provocation with threats and counterthreats that were calculated to deter the opponent at optimal levels, de-escalate tensions, and promote crisis stability. Although leaders may have tried to keep their beliefs, attitudes, perceptions, feelings, and actions mutually consistent in order to deal with stress during their disputes (an exceedingly difficult practice for critics of rational choice to demonstrate), there was no evidence that this adversely affected the choice of strategy. It may be true that a complete understanding of crucial foreign policy decisions cannot be obtained without reference to leaders' previously established beliefs, but these beliefs did not appear to have a noticeable impact on the crisis-management techniques employed by either side. In fact, notwithstanding differences in culture, ideology, political system, religion, history, leadership skills, belief structures, idiosyncrasies, and so on, there was an apparent uniformity of approach, a strategic rationale that seemed to guide superpower relations during foreign policy crises. Moreover, these patterns were present despite changes in the distribution of nuclear capabilities – from U.S. strategic dominance in the '40s and '50s to approximate parity in the late 1960s.

The distinct patterns emerging from analyses of superpower rivalry may also explain the discrepancies between rational choice (game-theoretic) and political psychology schools. If nuclear rivals are more prone to exhibit patterns of rational choice and restraint, and less apt to succumb to potentially harmful decision pathologies, then exclusion of these cases of superpower rivalry biases results against support for rational deterrence theory. Jervis, Lebow, and Stein's extensive work (1985) on crises in the Middle East, for instance, emphasizes decision making within Israel and Egypt, but U.S.-Soviet interaction is generally excluded from detailed analysis.⁵ If the level of threat experienced by officials in Egypt and Israel was not sufficiently potent to stabilize hostilities, then perhaps there is an important distinction that should be drawn between nuclear and conventional deterrence that is rarely acknowledged by critics.

Some have claimed that nuclear weapons were irrelevant to the long post-World War II peace, and "neither crucially define[d] a fundamental stability nor threaten[ed] severely to disturb it" (Mueller 1988, 55). The stable relationship among major powers since 1945, according to proponents of this position, would have developed regardless of the bomb, because it was the horror associated with even a conventional battle that served as the real deterrent. They have argued further that

the utility of nuclear weapons simply cannot be measured in the absence of nuclear war; there is no way to demonstrate that intercontinental ballistic missiles (ICBMs) prevented the Soviets from invading Western Europe or deterred the United States from attacking during the Cuban missile crisis. Consequently, the assertion that nuclear weapons were, and remain, irrelevant is as valid and defensible as the alternative claim that they did produce a stable superpower relationship. On the other hand, if it is possible to assess the role of these weapons in the absence of a nuclear war by moving beyond the restrictions of the dominant testing strategy and by focusing on a much wider range of propositions related to game theory, nuclear rivalry, and crisis escalation (as presented in chapters 4 and 5), then evidence produced by this type of analysis should be considered prior to making final judgments about the relative utility or disutility of nuclear weapons. To claim that there is but one way to test the many propositions underlying nuclear or conventional deterrence would amount to nothing less than premature closure of inquiry.

*Nuclear Proliferation in the Twenty-First Century:
Optimism and Pessimism Revisited*

Ordinarily, demonstrating that U.S.-Soviet crisis behaviour was cautiously restrained should elicit feelings of relief. After all, empirical support for OT and GRIT implies that the nuclear-based bipolar system, with all its faults, created an environment that not only prevented direct hostilities between the United States and the USSR but appeared to facilitate the management and prevention of prolonged, violent hostilities between client states within each sphere, as Waltz (1964) predicted. However, if the stability of the Cold War was a product of the rationality and cautious restraint associated with nuclear rivalry, then the transformations the system has experienced and is currently undergoing may produce a nuclear environment substantially less stable than the one left behind, notwithstanding the apparent short-term advantages to a world free of U.S.-Soviet tension.

Of course, "optimists" have argued that horizontal proliferation of nuclear weapons could actually contribute to crisis stabilization (Waltz 1981, 1985); restraint in bargaining and compromise would be expected to occur more frequently, all other things being equal. There is some evidence that states engaged in conflicts not involving a nuclear restraint experience a higher probability of war (Buono de Mesquita and Riker 1982). Others have defended the potential stabilizing effects of controlled proliferation in Europe and the Third World with recommendations calling for a "well-managed" spread to

compensate for the destabilizing effects of asymmetric distribution (Mearsheimer 1990a). Similar arguments have been used to support maintenance of a Ukrainian nuclear program (Mearsheimer 1993) and to defend proliferation in the Middle East to control Arab-Israeli conflict (Feldman 1981; Rosen 1977). Perhaps Waltz (1993, 554) is right: the spread of nuclear weapons “is something we’ve worried far too much about and tried too hard to stop.” The measured spread of nuclear weapons might be something to welcome, not fear.

The problem with this line of reasoning has been dealt with in detail elsewhere, so only a brief review of the “pessimist” response will be covered here, leading to a discussion of policy recommendations later on (chap. 7).⁶ First, the acquisition of nuclear weapons by non-nuclear powers does not *in and of itself* satisfy the conditions for stable nuclear deterrence. The optimists’ predictions would hold only if new nuclear powers were able to deploy their forces in a manner that satisfied a certain set of requirements, many of which were stipulated by Wohlstetter in 1959:

- 1 They must maintain a standing army and a reliable deterrent force in peacetime.
- 2 This force must be capable of surviving a preemptive first strike.
- 3 National leaders must be able to make the decision to retaliate and then transmit this command to the military forces; that is, there must be a reliable command, control, communication, and intelligence network – C3I.
- 4 New nuclear powers must have the ability to penetrate active defences of the enemy anti-aircraft missiles and interceptors, for example.
- 5 They must have the ability to overcome passive defences of the enemy bomb shelters and other civil defence capabilities, for example.

If deployment fails to meet these criteria, then nuclear deterrence, which traditionally has depended on the ability to threaten retaliation with a credible and effective second strike, would not apply. In fact, preemption in a crisis involving two new nuclear rivals would become more likely, given the incentives to launch on warning to avoid suffering overwhelming losses in a first strike. Although U.S.-Soviet nuclear deterrence functioned according to theory, it is doubtful that horizontal proliferation would contribute to crisis stabilization in every case. In addition to the weapons themselves, new nuclear states would need a sophisticated C3I network (item 3 in the list), an essential component of the relatively stable superpower relationship. Until such systems are developed to accommodate prospective nuclear rivalries, the

post-Cold War system is likely to become less stable as a consequence of proliferation.

Even u.s. and Soviet officials were not immune from communication and control problems during their crises, especially in cases where forces were placed on nuclear alert status (Sagan 1985). Trying to coordinate political and military objectives through these costly signals can undermine even the best deterrent efforts. Sagan (1994, 66) has recently explored several organizational impediments that affect crisis management, arguing that "common biases, inflexible routines, and parochial interests" often make military organizations prone to behaviours that lead to deterrence failures. The only safeguard is to create "tight and sustained civilian control," a strategy that leaders of new nuclear states are unlikely to embrace. Finally, even if one were to accept the optimists' prediction that proliferation would prevent wars, peace under those conditions might not always be preferred. War would not have been an option, for example, if the Gulf War coalition had been facing an Iraq that had nuclear arms. Peace would have prevailed, but at what cost?⁷

Pessimists have built a very compelling case for the conclusion that the only practical solution is to continue strengthening nonproliferation regimes by placing strict limits on the distribution of all nuclear materials. The less appealing alternative would be to insure that countries with nuclear weapons, or those in the process of acquiring them, are provided with the right kind of weapons technology (C31) to meet the conditions for a stable and effective deterrent retaliatory threat. The question is whether this second alternative really is less appealing.

In a comprehensive study of proliferation in the aftermath of the Soviet collapse, Ellsberg has documented the potential for "haemorrhaging of fissionable material" from stockpiles of plutonium in the former Soviet nuclear complex. As he explains, "the breakdown of central authority and military discipline in the Soviet Union has created a situation in which practical control over the Soviets' 27,000 warheads – and especially the 17,000 tactical warheads – is very much in question" (1992, 139–41). The assurances by President Yeltsin that these weapons will remain under a centralized command are not sufficient to eliminate all uncertainties. Even if Yeltsin is successful in his efforts to establish a unified command structure, there are three thousand to five thousand experts with top security clearance (out of approximately ten thousand) that have either lost, or experienced drastic cuts in, their incomes (140). These atomic scientists are under immense pressure to provide, for hard cash, the expertise, and possibly materials, needed to develop nuclear programs.⁸ The average wage for an American or Japanese scientist is roughly u.s.\$100,000 a year,

compared to the current rate for their Soviet counterparts of \$500 a month, and \$100 a month for a skilled technician (Clancy and Seitz 1992). Notwithstanding the enormous financial commitment by the U.S. and European governments to establish a number of thinktanks throughout the former Soviet Union to keep these experts from leaving, many scientists will no doubt be anxious to sign lucrative contracts in other countries that far exceed the economic incentives to stay.

Confirmation of the advanced stage of Iraq's nuclear program in the aftermath of the Gulf War clearly illustrated the ineffectiveness of the International Atomic Energy Agency (IAEA) and other monitoring agencies.⁹ More recently, North Korean officials refused UN inspection of two nuclear facilities and threatened to withdraw from the Non-Proliferation Treaty. One can only speculate about the effects this will have on decisions in South Korea, Japan, and Taiwan regarding their adherence to the treaty. And, although technically under Russian control, Ukraine continually refuses to sign a legally binding agreement to hand over all former Soviet nuclear weapons. As more states begin to acquire the requisite material, others will reassess the utility of the nonproliferation option given their own security interests. Nuclear regimes have managed to prolong the time lag between acquiring the material and developing weapons, but they have not been very successful at stopping the spread. As one senior official of the IAEA observed "proliferation has already happened. The main problem ... is not so much preventing the spread of nuclear weapons, but making it survivable" (*Time* 1985, 36), a position that appears to be consistent with the relatively moderate U.S. response to North Korea's decision.

The only concrete proposals that continue to receive official support are those that recommend repeating empty promises to re-sign an outdated treaty. The following warning, offered by Clancy and Seitz in the conclusion to their excellent, though highly technical, study of nuclear proliferation and technology transfers after the Cold War, is particularly appropriate here: "Contemplating the realm of science and technology today, and considering the prospect of what intelligence, human and artificial, can accomplish in the century to come, we arrive at the depressing conclusion that the present regime of proliferation control may be among that rare set of entities – a fit object for catastrophe theory" (1992, 12).

It is time to reconsider proposals that call for a more managed or controlled distribution of CGI technology in order to reduce the destabilizing effects of asymmetric proliferation.¹⁰ Given the fast pace

of change following the collapse of the Soviet Union, and the overwhelming incentives to spread and acquire the necessary technology, experts and policymakers should give these less appealing alternatives serious consideration. Although counterintuitive, they are at least not entirely dependent for their success on our blind faith in international law. Even pessimists like Sagan (1994, 106) agree that some effort should be made to help new nuclear states acquire the right kind of technology and that the United States should work towards making new nuclear forces survivable and invulnerable by "cooperating with new proliferators [and] sharing information on delivery systems technology, operational practices, and advanced warning systems."¹¹ In the long run, a well-managed distribution of C3I technology could generate the kind of restraint and stability characteristic of the U.S.-Soviet relationship. It is true that even their relationship was dangerous on occasion, but the real choice that should be guiding consideration of policy alternatives today is not between living in a world with nuclear weapons and a world without them, but between living in a world with many nuclear states that is stable and one that is highly unstable.

CONTEMPORARY RELEVANCE OF U.S.-SOVIET RIVALRY

In view of the claims in the literature that the analysis of U.S.-Soviet crisis behaviour holds merely antiquarian interest for today's less confrontational world, I offer a few thoughts on why insights gained from studying their rivalry remain relevant.

First, it should not be assumed that the old nuclear rivals are capable of preventing or even controlling events throughout the world that might cause them to experience mutual crises. States such as China, India, Pakistan, Ukraine, Syria, and Israel remain capable of generating intense international conflict, independently of American or Russian wishes. Second, the two traditional rivals must cope with a world of multiple nuclear powers in which some disputes not initially involving the superpowers could escalate to the level of nuclear confrontation. So, even if superpower rivalry is over, there is still value in assessing the record of U.S.-Soviet crises. Given the likely proliferation of nuclear weapons over the long term, the discovery of patterns of interaction should facilitate understanding of nuclear rivalry more generally. The inference that little of value has been gained from forty years of research on U.S.-Soviet crises and nuclear deterrence and that a radically new approach is required to assess the implications of change today is a bit premature. Efforts towards greater understanding

of the prospects for nuclear stability, based upon detailed examination of the longest nuclear rivalry in history, not only are valuable but have become imperative.

Other factors further undermine the assumption that reduced direct confrontation between the superpowers will necessarily result in harmonious relations at a more general level. It would be extremely optimistic to assume that reforms in Europe somehow will overwhelm the legacy of generations of conflict in the Third World and other parts of Europe. Although the United States and other Western powers continue to have a very strong interest in preventing such crises, many states, including the United States, are beginning to respond to domestic pressures to turn inward, now that the Cold War is over. Events in Yugoslavia have demonstrated that Western indifference to international disputes could prove to be disastrous.

The process of change itself provides yet another reason to suspect that existing nuclear powers still could end up in crises. With more independent governments functioning in Eastern Europe, it is not at all clear that the United States and Russia will find the management of global nuclear security any easier. Internal problems within Russia or Ukraine could make it difficult for Yeltsin or his successor to pursue a conciliatory foreign policy. Although a change in leadership seems very unlikely at this time, conservative elements in Russia may be joined by others in demanding a more assertive foreign policy, as occurred in the former Yugoslavia with Russia's initial Security Council veto over economic sanctions on Serbia, as Russia became more assertive in its defense of Serbia against the United States, NATO, and the UN.

7 Conventional Deterrence and Compellence Theory: Perspectives on Testing after the Cold War

The focus of this final chapter is conventional deterrence and compellence theory. The purpose is to develop an argument in favour of an alternative testing strategy in this realm using protracted crises as the main source of empirical evidence. The chapter unfolds in four stages. Stage one briefly summarizes the key impediments to testing derived from chapter 2. Stage two introduces a different approach that recommends identifying separate deterrence and compellence encounters within a single foreign policy crisis, thus expanding the pool of evidence that would be appropriate for testing a wide range of theoretical propositions derived from theory. Stage three describes, in summary form, fourteen immediate deterrence/compellence exchanges between officials of the United States, NATO, and the United Nations and Bosnian Serb leaders that took place from April 1993 to September 1995; it considers whether the prerequisites for effective use of these coercive strategies were met, and it assesses whether the behaviour in these encounters was consistent with the theoretical predictions. Stage four addresses policy implications and the overall contributions of a protracted crisis approach to testing.

As described in chapter 2, the most prominent strategy used to produce evidence to evaluate rational deterrence theory recommends identifying cases of immediate deterrence, coding these cases as instances of success or failure, isolating conditions that were present (absent) during successes and absent (present) during failures and, based on these differences, drawing conclusions about why and how deterrence works. In the empirical domain, lack of correspondence

in case selection and coding is the key area of difficulty confronting those who apply this success/failure framework. The following are among the many coding questions that must be answered before making valid judgments about deterrence:

- 1 Who is the challenger and who is the defender in each case? Since military-security crises involve a series of interactions and deterrence episodes, with each side, and their respective clients, acquiring and playing both roles at various stages, disagreements about who initiated the crisis are common.
- 2 Is the retaliatory threat direct or extended?
- 3 Is the crisis a deterrence or compellence encounter or some combination?
- 4 What class of deterrent or compellent threat is being issued?
- 5 Finally, does the case constitute a success or failure or some combination of both?

Case selection is likely to be difficult for researchers who choose to test deterrence theory this way, given the many opportunities available to reject any one case. As revealed by the ongoing debates over case listings described in chapter 2, these obstacles cannot be overcome through reference to the historical record; in fact, each side has offered compelling evidence to support their distinct, and in some cases contradictory, interpretation of events in almost every crisis. Once again, while debates over the accuracy of historical accounts are constructive, lingering divisions become counterproductive if very little effort is directed towards producing different testing strategies that lie outside the success/failure framework, alternative sources of empirical evidence and data, or a wider range of propositions derived from the theories.

This chapter offers an alternative method of testing deterrence and compellence theory that avoids at least some, although not all, of the coding controversies noted in chapter 2. It does this by rejecting the assumption that a crisis encompasses within it a single, dominant encounter. Instead, each case is viewed as a series of separate and distinct deterrence and compellence exchanges, thus expanding the pool of evidence available to test the theories. Dissecting each crisis to reveal different encounters allows for multiple interpretations of any one foreign policy crisis and, more importantly, can account for discrepancies across existing case lists. The approach is expected to provide a fairer test of the theory by specifying the precise time frame and exact sequence within which the appropriately designated threats, counterthreats and responses were made. Although some scholars try

to deal with this problem by partially specifying the class of deterrence they are concerned with (namely, threats of military retaliation to deter military attacks), the potential for confusion remains. Regardless of efforts to focus on features of a crisis that appear, on the surface, to be an immediate deterrence encounter, the behavioural properties of other classes of deterrence and compellence may influence actions and outcomes.

Others have attempted to solve the problem of selection bias by expanding the empirical domain of deterrence theory in a slightly different way. Instead of focusing on individual cases of deterrence failure or breaking a crisis down into a series of discrete exchanges (as proposed here), they claim that a valid test requires an assessment of reputation and its impact on perceptions of capability, commitment, and resolve within enduring rivalries (Lieberman 1995; Geertz 1995). Based on a study of deterrence relationships in the Middle East over time, Lieberman concludes that Egyptian and Israeli behaviour supports the theory's predictions – long-term deterrence stability (success) is achieved through short-term failures, because only failures can provide an opportunity for leaders in the defending state to demonstrate resolve, capability, and credibility, all essential components of successful deterrence. Evaluating deterrence in the context of rivalries helps to explain otherwise puzzling phenomena and provides a new way to test old hypotheses. It does this by making the entire dispute the unit of analysis, so the relationship between success and failure can be probed for information about the dynamic nature of deterrence unaccounted for in traditional approaches (Geertz 1995).

On the other hand, an enduring-rivalry approach is not sufficient to address the most important puzzle that continues to plague the research program on deterrence, namely the discrepancies across case lists. In the absence of consensus on coding individual cases within any given rivalry, researchers will remain ill-equipped to address the new and interesting questions about the relationship between success and failure that a rivalry approach offers. In fact, they run the risk of spinning their wheels in their efforts to move the rational deterrence debate forward.

A protracted-crisis approach is important for two reasons: it accounts for anomalies across current data sets, as previously described, and it can provide answers to new questions about deterrence in rivalries. The difference is that the dynamic nature of the deterrence relationship is explored in the context of a single, protracted crisis. The distinction between the three testing strategies is set out in Table 7.1. The effects of learning, reputation, and background on resolve, credibility, and deterrence stability all remain part

Table 7.1

Empirical Domain of Immediate Deterrence/Compellence Tests (Coding Rule:
A foreign policy crisis represents...)

<i>Multiple Exchanges in Protracted Crises</i>	<i>A Single Dominant Exchange Viewed Independently</i>	<i>One Exchange in Protracted Rivalry</i>
This study George and Smoke (1974)	Huth and Russett (1984, 1990, 1993) Lebow and Stein (1989a, b, 1990)	Lieberman (1994, 1995) Geertz (1996)

of the research program, thus providing a strong empirical base for evaluating theory.

DETERRENCE AND COMPELLENCE
IN PROTRACTED CRISES:
BOSNIA-HERZEGOVINA, 1993-95

Definitions and Coding

Huth and Russett's definitions of deterrence and compellence are used to distinguish the two categories of coercive threats. "Compellence" is defined as "an attempt by policymakers in state A to force, by threat and/or application of sanctions, the policy of state B to comply with the demands of state A, including but not limited to, retracting actions already taken." Deterrence occurs when "the threatened sanction is designed to prevent state B from taking actions it is considering but has not already initiated; thus the sanction would be employed only if the target undertook the action that the deterrer had sought to prevent" (1990). Since the objective in each case is to prevent undesired actions, either by the threat of sanctions or the offer of rewards, both forms of coercive diplomacy can be viewed as essentially similar strategies – preventing undesired actions can take the form of compelling a state to retract or discontinue actions already taken, or deterring a state from taking actions it has contemplated but not initiated.¹

The relevant question is whether U.S., European (NATO), and UN encounters with the Bosnian Serbs between 1993 and 1995 constituted specific instances of immediate deterrence or compellence. If retaliatory threats were not initiated, the exchange would be inappropriate for evaluating theory. For it to qualify as an immediate deterrence or compellence exchange, therefore, the Bosnian Serbs must have been considering (or already undertaking) an action that was viewed by their opponents as undesirable and, in response, an attempt must have been made to dissuade the Bosnian challenger from committing the undesired

action, through the use of a threat of sanctions, an offer of rewards or inducements, or both.² Once a specific encounter is deemed relevant, a judgment is made about success or failure based on the following prediction derived from the theory: a retaliatory threat will succeed if (A) leaders define the unacceptable behaviour and communicate to challengers a commitment to punish violations; (B) the threatened punishment is severe enough to deny the challenger the objectives sought;³ (C) the deterring state possesses the capability to do so; and (D) the leaders demonstrate their resolve to carry through with the threat (Lebow and Stein 1990). Resolve is most effectively demonstrated through costly signals – that is, through any action, statement, or condition that increases the political, economic, or military costs assigned to the status quo, while lowering the costs of responding to a challenger's probes. These demonstrations usually take the form of one or more of the following:

- 1 actions – for example, deployment of air, sea, or ground forces; evacuation of peacekeepers from safe havens, thus allowing a more decisive air strike response to probes;
- 2 statements – for example, public announcements (promises) of impending retaliation; explicit ultimatums and deadlines; public displays of unity among coalition members in support of response;
- 3 domestic support for retaliation – for example, supportive public opinion; positive domestic and international press coverage.

If these conditions (A–D) are satisfied, the expected net costs, to the challenger, of the threatened sanction should be greater than the expected net gain of noncompliance, because the punishment, if carried out, would prevent the challenger from achieving intended goals. If these requirements are met, but the behaviour still occurs, that would constitute a failure – both in theory and strategy. On the other hand, if one or more of the conditions is not satisfied, the theory predicts failure in most cases. In other words, even clear and credible threats will fail if the challenger believes that the challenge is worth the risks and political, military, or economic costs incurred by triggering the threatened response.⁴

Data and Methodology

All major statements, threats (implicit or explicit), and actions (sanctions, mobilization of force, demonstration of force, dispatch of diplomats, and so on) initiated by either U.S., NATO, or UN leaders to alter (deter or compel) the behaviour of the Bosnian Serbs were recorded from April 1993 through September 1995.⁵ All major political

and military responses to those threats were identified, and special attention was given to the time between threat and response (minutes, hours, or days). This approach is crucial when judging the success or failure of coercive threats, and the failure to follow it accounts, once again, for discrepancies across other case lists. Four main sources of data were used: interviews with personnel in defense policy and planning divisions within NATO; unclassified UN and NATO documents; *Keesing's Contemporary Archives*; and the *New York Times* and *New York Times Index*.⁶ Detailed coverage of the crisis in the *New York Times* and *Keesing's* provides an excellent chronology of events and interactions for the period in question. This data also facilitates assessment of the nature, sequence, and timing of provocations and threats of retaliation throughout the year. In order to deal with the problem of biased reporting and the possibility that some judgments about the clarity, credibility, and capability of U.S./NATO threats are entirely impressionistic, two NATO officials and Major-General J.A. MacInnis (United Nations Protection Force) had an opportunity to review the evidence and offer their interpretations of perceptions, intentions, and motives in each exchange.⁷

Summary and Findings

In almost every encounter over the two-year period, coercive diplomacy succeeded – and failed – for all the right reasons. A summary of interactions between 1993 and 1995 is presented in Table 7.2.⁸ The column headings (numbers 1 to 14) represent separate exchanges. Exchanges with a/b designations represent encounters with more than one defender and/or challenger, or exchanges in which the defender and challenger changed roles. The letters Y and N (for Yes and No) represent, respectively, the presence and absence of key prerequisites (A–D) for deterrence or compellence success, as stipulated in the discussion of definitions and coding earlier in this chapter. (But in the final row, Y and N represent the outcome, success and failure, respectively.) Exchanges with split codings (for example, Y/N) represent cases in which two different threats were initiated during the same interaction. It goes without saying that complete agreement on these coding decisions is not likely; some will argue that certain exchanges should be merged (or separated further), while others may feel that conditions coded as present (Y) were really absent (or vice versa). The crucial point is that traditional debates over the coding of entire crises (in which both sides may be correct, depending on the time frame) are replaced by disputes over more specific threat-response interactions that force researchers to track the sequence of events in more precise

Table 7.2
 Summary of Exchanges with Bosnian Serbs, 1993–95

<i>Prerequisites</i>	<i>1</i>	<i>2a/b</i>	<i>3a/b</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>8</i>	<i>9</i>	<i>10</i>	<i>11</i>	<i>12</i>	<i>13a/b</i>	<i>14a/b</i>
Unacceptable behaviour defined and threat communicated	Y	Y/Y	Y/Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y/Y	Y/Y
Threat potentially costly to challenger	Y	Y/N	Y/N	Y	Y	N	Y	N	Y	N	Y	Y	Y/Y	Y/Y
Deterring state has capability to deny gains from challenge	Y	Y/N	Y/Y	N (light option)	Y	Y	Y	N	Y	Y	Y	Y	Y/Y	Y/Y
Resolve (consensus) demonstrated	Y	Y/N	Y/N	N	Y	N	Y	N	N	N	N	N	N/Y	Y/Y
Success?	Y	Y/N	Y/N	N	Y	N	Y	N	N	N	N	N	N/Y	N/Y

terms. Consequently, far more relevant information is made available upon which to judge the strengths and weaknesses of deterrence and compellence theory, since we are no longer faced with having to fit all of this information into a single data point. Even if some of the coding decisions in Table 7.2 are incorrect (even if, say, only three exchanges are coded accurately), that would be sufficient evidence in favour of probing existing case lists for more data points.

The approach also lends itself well to aggregate testing. Tables 7.3 and 7.4 offer one illustration of how we might evaluate deterrence theory more fairly, based on information obtained from expanding current data sets. In order to set the stage for the argument, assume we have complete agreement on each of five coding decisions (four prerequisites and the outcome) for every case in a hypothetical data set of 180 deterrence exchanges – a difficult starting point, to say the least. Assume also that the four prerequisites have a uniform impact on success/failure and that outcome is a dichotomous variable (Y/N), two issues we will return to later. Viewed from the perspective of alternative scenarios, or response sets, there are sixteen combinations that are consistent (c) with predictions derived from rational deterrence theory (Table 7.3) and sixteen combinations that are inconsistent (i) with expectations (Table 7.4). Moreover, some response sets offer stronger confirming (w_A)/disconfirming (w_B) evidence than others, depending on the number of prerequisites satisfied and the outcome produced. With respect to scenarios consistent with predictions, for example, cases that conform to response set 1c, in which all four conditions are present and the deterrent threat succeeds (Y), should count as stronger confirmation of the theory than all others, with the exception of 16c – it too receives a weight of four, given theoretical expectations. Similarly, response sets 12c–15c, in which only one prerequisite is satisfied and fails to deter an opponent, should have more confirming weight than those in which three conditions are satisfied yet fail to prevent a challenge (for example, 2c–5c). The same logic was used to assign strong and weak “disconfirming” weights to inconsistent scenarios, w_B and $(4 - w_B)$ respectively. Based on the distributions of cases across these response sets, and by comparing strong/weak confirmations to strong/weak disconfirmations, it is possible to calculate a rudimentary coefficient of predictability, or reliability, for deterrence theory.

$$(1) \frac{\sum n(w_A)}{\sum n(w_A) + \sum n(4 - w_A)} = 65.0\%$$

Table 7.3
 Consistent Response Sets (Assumes prerequisites have equal weight)

Response Sets	<i>Deterrence/Compellence Prerequisites</i>				Success?	w_A	$(4 - w_A)$	n	$n(w_A)$	$n(4 - w_A)$
	A	B	C	D						
1c	Y	Y	Y	Y	Y	4	0	18	72	0
2c	Y	Y	Y	N	N	1	3	4	4	12
3c	Y	Y	N	Y	N	1	3	3	3	9
4c	Y	N	Y	Y	N	1	3	9	9	27
5c	N	Y	Y	Y	N	1	3	7	7	21
6c	Y	Y	N	N	N	2	2	11	22	22
7c	Y	N	N	Y	N	2	2	1	2	2
8c	N	N	Y	Y	N	2	2	5	10	10
9c	N	Y	N	Y	N	2	2	7	14	14
10c	Y	N	Y	N	N	2	2	1	2	2
11c	N	Y	Y	N	N	2	2	3	6	6
12c	Y	N	N	N	N	3	1	4	12	4
13c	N	N	N	Y	N	3	1	2	6	2
14c	N	N	Y	N	N	3	1	10	30	10
15c	N	Y	N	N	N	3	1	10	30	10
16c	N	N	N	N	N	4	0	13	52	0
Total								108	281	151

Table 7.4
 Inconsistent Response Sets (Assumes prerequisites have equal weight)

<i>Response Sets</i>	<i>Deterrence/Compellence Prerequisites</i>				<i>Success?</i>	w_n	$(4 - w_n)$	n	$n(w_n)$	$n(4 - w_n)$
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>						
1i	Y	Y	Y	Y	N	4	0	12	48	0
2i	Y	Y	Y	N	Y	1	3	7	7	21
3i	Y	Y	N	Y	Y	1	3	2	2	6
4i	Y	N	Y	Y	Y	1	3	5	5	15
5i	N	Y	Y	Y	Y	1	3	0	1	0
6i	Y	Y	N	N	Y	2	2	7	14	14
7i	Y	N	N	Y	Y	2	2	0	0	0
8i	N	N	Y	Y	Y	2	2	4	8	8
9i	N	Y	Y	N	Y	2	2	9	18	18
10i	N	Y	N	Y	Y	2	2	5	10	10
11i	Y	N	Y	N	Y	2	2	0	0	0
12i	Y	N	N	N	Y	3	1	2	6	2
13i	N	N	N	Y	Y	3	1	12	36	12
14i	N	N	Y	N	Y	3	1	3	9	3
15i	N	Y	N	N	Y	3	1	1	4	1
16i	N	N	N	N	Y	4	0	3	12	0
Total								72	180	104

$$(2) \frac{\sum n(w_A) + \sum n(4 - w_B)}{(\sum n(w_A) + \sum n(4 - w_B)) + (\sum n(w_B) + \sum n(4 - w_A))} = 53.7\%$$

Equation 1 represents the proportion of strong (w_A) confirmations to weak ($4 - w_A$) confirmations across consistent response sets, while equation 2 offers a more comprehensive assessment of the theory's predictability by combining all information about strong and weak confirmations/disconfirmations across both categories, which explains the lower predictability score. When applied to the 14 exchanges in Bosnia (Table 7.5), the results are 65.6 percent and 61.7 percent for equations 1 and 2, respectively.

$$(1) \frac{\sum n(w_A)}{\sum n(w_A) + \sum n(4 - w_A)} = 65.6\%$$

$$(2) \frac{\sum n(w_A) + \sum n(4 - w_B)}{(\sum n(w_A) + \sum n(4 - w_B)) + (\sum n(w_B) + \sum n(4 - w_A))} = 61.7\%$$

These results are encouraging for proponents of deterrence theory.

All of this assumes, of course, that the four prerequisites are equally important in determining success and failure. If they vary with respect to relative potency, then an additional level of weighting should be assigned to each response set to provide a more accurate representation of confirmations and disconfirmations of the theory. If resolve (condition D) is more important, for example, a higher confirmation weight would be assigned to response set 2c (for example, $w_A = 4$, instead of 1), and the percentage values for equations 1 and 2 would be significantly higher, thus representing even stronger confirmation of the theory. On the other hand, resolve may be less important, for instance, in situations where the defender's capabilities are so overwhelming and the costs of retaliation are so low that the deterrent threat remains credible even if resolve is questioned. It is also conceivable that the relative importance of the prerequisites varies from crisis to crisis, and perhaps even from exchange to exchange, depending on circumstances. For example, throughout most of the first year of fighting in Bosnia-Herzegovina, NATO's capability (condition C) to deny the Bosnian Serbs the objectives they sought was never really doubted, until exchange 8. It was during this exchange that the Bosnian Serb leadership began to question the ability of UN and NATO

Table 7.5
 Consistent Response Sets: Bosnia, 1993–95 (Assumes prerequisites have equal weight)

<i>Response Sets</i>	<i>Deterrence/Compellence Prerequisites</i>				<i>Success?</i>	w_i	$(4 - w_i)$	n	$n(w_i)$	$n(4 - w_i)$
	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>						
1c	Y	Y	Y	Y	Y	4	0	7	28	0
2c	Y	Y	Y	N	N	1	3	4	4	12
3c	Y	Y	N	Y	N	1	3	0	0	0
4c	Y	N	Y	Y	N	1	3	0	0	0
5c	N	Y	Y	Y	N	1	3	0	0	0
6c	Y	Y	N	N	N	2	2	1	2	2
7c	Y	N	N	Y	N	2	2	0	0	0
8c	N	N	Y	Y	N	2	2	0	0	0
9c	N	Y	N	Y	N	2	2	0	0	0
10c	Y	N	Y	N	N	2	2	3	6	6
11c	N	Y	Y	N	N	2	2	0	0	0
12c	Y	N	N	N	N	3	1	2	6	2
13c	N	N	N	Y	N	3	1	0	0	0
14c	N	N	Y	N	N	3	1	0	0	0
15c	N	Y	N	N	N	3	1	0	0	0
16c	N	N	N	N	N	4	0	0	0	0
Total								17	42	22

Note: 1c (1, 2a, 3a, 5, 7, 13b, 14b); 2c (9, 11, 12, 13a); 6c (4); 10c (3b, 6, 10); 12c (2b, 8); see Table 7.2.

officials to coordinate operations for air support. Several U.S. warplanes preparing to retaliate against a Serbian attack on French peacekeepers during the first week of April 1994 were unable to respond in time due to problems with communication in the chain of command, a product of the dual-key system for UN-NATO operations in the region. The original request for air strikes was made by the French commander whose troops were being attacked. His request went to the commander of peacekeeping forces in Bosnia, then to the commander of UN forces in Yugoslavia, and finally, two hours later, to Yasushi Akashi, the special representative of Secretary General Boutros Boutros-Ghali, who was the only person with the power to authorize the strikes. The UN official then tried, for approximately one hour, to contact Bosnian Serb military leaders to inform them of the impending strikes should they continue attacking the French troops, and only then requested that NATO planes respond. The three-hour delay allowed Serb forces and weaponry to escape without having to face air strikes. From that point onwards, NATO officials were faced with having to satisfy the capability prerequisite in order to mount a credible deterrent threat, whereas previously it was only NATO resolve that posed the most difficult problem for peacemaking efforts.

In addition to the question of variations in relative potency, two other questions need to be addressed before making final judgments about deterrence and compellence: What are the interrelationships among the four prerequisites, and how do they affect relative importance and weighting? and How can we develop a more accurate operationalization of success and failure that accounts for partial defeats and victories? Constraints on space preclude a more detailed treatment of these central issues. The purpose of raising the questions here is to reemphasize Huth and Russett's (1990) observation that "rigor makes a difference" when testing deterrence. The theory continues to suffer from premature closure of inquiry because of the prevailing assumption that there is only one valid way to test. It is interesting to note that the most widely cited critique of deterrence (Lebow and Stein 1990) focuses almost exclusively on identifying crises that fit two of the thirty-two combinations noted in Table 7.3, namely, response set 1c and 1i. Although this approach is helpful, it excludes thirty other scenarios that provide equally crucial confirming and disconfirming evidence on which to evaluate the overall strength of the theory. The remaining thirty sets should be explored before making premature judgments about the utility of deterrence or compellence theory.

The data generated from this approach can also help identify necessary and sufficient conditions for deterrence success and, by implication,

the relative potencies of individual components, a fundamental prerequisite for policy recommendations.⁹ In the context of an immediate deterrence encounter, the theory stipulates that the presence of all four conditions will lead to success, and the absence of one or more will lead to failure. Working through the logic of necessary and sufficient conditions, we can produce several additional pieces of information about deterrence. For example, if at least one of the conditions is absent every time deterrence fails, the specified condition is sufficient for failure. If deterrence fails every time any one component is absent, then they are all necessary for success. Similarly, if the four prerequisites are present when deterrence fails (Lebow and Stein), we can conclude that none of the conditions individually is sufficient for success and that the conditions taken together are not sufficient for success. The presence of all four conditions every time deterrence succeeds provides no information about whether their presence is sufficient for success. Air, dirt, guns, and people are present at all deterrence successes, but their presence in no way ensures that outcome. One would also have to determine whether they are present whenever deterrence fails.¹⁰ If the four conditions are absent when deterrence succeeds (that is, the pattern is NNNNS), then the conditions are not necessary for success. Of course, if the defender does nothing to stop a challenge, and the challenger nothing to take advantage of the defender's inaction, then the exchange does not satisfy a defining conditions for an immediate deterrence encounter.

The preceding discussion also points to problems with applying the necessary and sufficient condition framework to deterrence – it assumes the four prerequisites and outcome are dichotomous variables (present/absent), rather than continuous. It also assumes the four components have the same weight with respect to outcome. Instead of viewing deterrence in absolute terms, a protracted-crisis approach evaluates the overall strength of the theory by assigning weights to inconsistent and consistent response sets. This avoids the tendency to reject theory simply by identifying anomalies that appear to be inconsistent with expectations derived from the necessary and sufficient condition hypothesis.

Many of the exchanges will appear to some to be open-ended and, in effect, overlapping. Critics will note that such artificially rigid periodization may render false assessments of success or failure, primarily because of emphasis on the short term. On the other hand, success and failure in this context is crucial for purposes of valid testing. If the Bosnian Serb leadership agreed to abide by UN or NATO demands in the short term, for instance, but probed for weaknesses and began to challenge the retaliatory threat over a period of time, that behaviour

should count as a failure of deterrence strategy, not deterrence theory. Western leaders lacked the political will to respond to probes by General Mladic in a way that would have demonstrated resolve. Since key components of the strategy were missing, and the retaliatory threats lacked credibility, the subsequent challenges by the Bosnian Serbs are consistent with expectations. In any case, a preference for viewing entire crises as single encounters, rather than as a series of distinct exchanges, would bias results in favour of identifying deterrence failures and would be no less prone to the risks of making false judgments about rational deterrence and compellence theory.

Finally, it is important to note that there will always be coding controversies, regardless of approach, because scholars make coding decisions based on their interpretation of the historical record. The approach recommended here will not solve all these problems. The more interesting question, however, is whether the new approach can help explain important anomalies and, at the same time, avoid producing new debates. Both of these tasks can be accomplished by identifying discrete exchanges, since debates over coding shift to a different, much more basic, level of analysis – namely, the sequence of specific threats and responses throughout a crisis.

THE ENDURING POLICY-RELEVANCE OF REALIST THEORIES

Policy-relevant theory may take on three distinct qualities (George 1991). It may be diagnostic, whereby emphasis is on describing how and why things work as they do. It may also take the form of a conditional generalization – that is, in situation *x*, if one does *y*, one should expect *z*. Finally, policy-relevant theory may be prescriptive, offering explicit recommendations to policymakers faced with certain kinds of problems. Deterrence and compellence theories (nuclear and conventional) encompass all these qualities. The evidence from chapters 4, 5, and 7 illustrates the utility of deterrence and compellence theories in both the diagnostic and conditional-generalization sense. Specifically, they provide useful frameworks within which to describe, explain, and predict events both during and after the Cold War. But the theories are also highly prescriptive in content, offering a set of specific policy guidelines to manage nuclear rivalry (as described in chapter 6) and control ethnic conflict. Among the less encouraging lessons of the Balkan war, for example, is that major powers are not likely to become seriously involved in preventing ethnic conflicts until it is clear that there are substantial political points to be gained. Since there is rarely any direct proof that diplomatic efforts accomplish

anything significant (even if they have), leaders are confronted with a particularly difficult task when trying to mount a large-scale effort at preventive diplomacy. The dilemma deterrence theorists face with respect to identifying successes, that is, proving that a retaliatory threat prevented a challenge, applies equally well to preventive diplomacy; it succeeds if nothing happens. The problem is obvious – if there is no war to stop, or any other concrete measure of diplomatic success, for that matter, leaders are not likely to make the first (often essential) move, notwithstanding the fact that early involvement may be cost-effective over the long run, especially in terms of lives saved.

The evidence suggests that deterrence and compellence theories, two abstract theories derived from analysis of interstate relations in anarchy, continue to make important contributions to the study of international conflict, crisis management, and security policy in a post-Cold War world. With all their faults, state-centric theories are as relevant to the study of ethnic conflict within states as they were to an understanding of U.S.-Soviet relations since 1945. The inference that nothing of value has been gained from the last fifty years of research on the causes and consequences of interstate war and ethnic conflict, that nothing of substance has emerged to facilitate an understanding of the properties of this form of violence and the prospects for stability, and that a radically new approach is required to assess the implications of change and strategies appropriate to deal with it is highly premature.

With respect to the contemporary relevance of state-centric theories like deterrence, it is important to note that the assumption that they are relevant will continue to be a central part of the research program on world politics, not only because of the greater theoretical clarity it provides but, more importantly, because the commitment to sovereign statehood remains a common objective for both old and new states (James 1993).¹¹ Writing more than thirty years ago, Arnold Wolfers (1959, 10) concluded that “Psychologically, nothing is more striking today than the way in which men in almost every part of the world come to value those possessions upon which independent national statehood depends ... [and] are willing to make the most sweeping sacrifices of their own well-being as private individuals in the interest of their nation.” As we approach the twenty-first century, commitments to “statism” have not diminished. On the contrary, the war in Yugoslavia, demands for sovereignty by the Baltic republics in the former Soviet Union, ethnic turmoil throughout Eastern Europe, and separatist movements from the former Yugoslavia to Canada are sufficient evidence that very little has changed in this regard. Furthermore, ethnic nationalist movements – often cited as increasingly important

nonstate actors – seek statehood, not transformation of the system. Analyses of the dynamics of international politics must take into account that a great deal of global activity, particularly with respect to questions of war and peace, but not limited to them, will continue to be a product of behaviour specified in the realist paradigm.¹²

Until there is sufficient evidence that an alternative set of assumptions is more useful, adherence to the realist paradigm and its attendant assumptions remains prudent. One should not assume that contemporary events are sufficient to render the contributions of an entire research program obsolete. Although it is true that complex interdependence has “undermined the conceptual unity of the state by perceiving it as an arena of competing bureaucratic entities,” and has partially “reduced the relative importance of the state by introducing a range of private transnational activity” (Cox 1986, 205), neither is sufficient to claim that realist theories are outdated. It is one thing to acknowledge growing levels of interdependence, but another to claim that we need new models and frameworks to understand contemporary international politics. Realism continues to provide a framework that facilitates development of international relations theory and foreign policy. Alternative frameworks may be very useful in explaining why conflict and war have become less likely today, but they provide almost no guidance for what to expect if and when conflict erupts. The possibility of military hostilities between major powers today may appear more remote, but this should not concern those studying the behaviour expected from antagonists who eventually will become involved in military-security crises in the future.¹³

The Postmodern Critique and Rejoinder

The postmodern critique emphasizes problems inherent in the methods of the realist paradigm and in its applications and research product over the past four decades.¹⁴ The claim is that international relations has not been cumulative, because of a fundamentally mistaken epistemology. Postmodernists simply reject the dominant scientific vision of reality that continues to provide the basis of knowledge construction and institutional design in the field. This position is hardly surprising, since postmodernism is an intellectual movement that is highly critical of any effort to develop theories of social action. It regards the dominant paradigm as flawed because it is based on a modernist proposition “which asserts that either we have some sort of ultimate ‘foundation’ for our knowledge or we are plunged into the void of the relative, the irrational, the arbitrary, the nihilistic” (George and Campbell 1990, 289).

The postmodern critique encompasses three distinct strands. All, however, are rooted in critical social theory and emphasize the inadequacy of the prevailing logical positivist or empiricist approach to knowledge construction. Some base their critique on language and linguistic determinants of reality (Phillips 1977; Giddens 1979; George and Campbell 1990); others stress problems stemming from the cultural bias of social research; and a final set of criticisms focuses on change and the time-specific, or historical, limitations of IR theory (George and Campbell 1990, 270). Each of the arguments will be evaluated in turn, with specific reference to recent contributions.¹⁵

Linguistic Relativism. One approach of postmodernists is to point to the complex nature of language and meaning as a critique of positivism; this critique is, in turn, relevant to the overwhelming amount of work in IR (Phillips 1977; Giddens 1979; George and Campbell 1990). Although a comprehensive assessment of the linguistic relativism debate is beyond the scope of this project, it is possible to address the underlying philosophical argument, which is fairly straightforward. Building on the work of Wittgenstein (1968), the linguistic variant of the criticism contends that any attempt to reduce everyday terms "to a singular essentialist meaning" is problematic given "the multiplicity of meaning to be found in social activity" (George and Campbell 1990, 273). By implication, a concept, term, word, or symbol cannot correspond "to some ... externally derived foundation or object" and ultimately is context-dependent. Similarly, Phillips argues that the validity of theory cannot be determined because "There is no standard or objective reality (always fixed, never changing) against which to compare a universe of discourse ... nothing exists outside of our language and actions which can be used to justify ... a statement's truth or falsity" (1977, 273).

Of course, it is not entirely clear how this "multiplicity of meaning" is sufficient to render meaningless an approach that assumes the existence of an objective reality. An important distinction must be drawn between the assertion that these discrepancies *might* have a significant impact on scientific theorizing and the assertion that they *do* have such an effect. In most cases, errors of interpretation and generalization produced by linguistic nuances are relatively insignificant and ultimately have very little impact on the generalizability of social theories. There are numerous words, symbols, concepts, and ideas, for example, that are commonly understood, regardless of other linguistic variations, but the implications of this standardized conceptual framework are frequently overlooked and ignored in the postmodern critique.

In any case, it is contingent upon the theorist to specify the precise meaning of any variable or symbol that is central to a theory. Although definitions may vary – possibly partly, but not entirely, as a consequence of language – scholars nevertheless are more likely than not to understand and agree on the underlying meaning of most words, symbols, and phrases. The point is that theorists generally do have a common starting point and often suspend, at least temporarily, counterproductive debates over meaning in order to shift emphasis towards the strength and logical consistency of the theory itself, a more important issue that has nothing to do with language. Evaluating the internal consistency of the central assumptions and propositions of a theory, that is, criticising from within, is likely to be more conducive to theoretical progress than the alternative, which is to reject the idea of theory building entirely.

Finally, the lack of purity and precision, another consequence of linguistic relativism, does not necessarily imply irrelevance of purpose or approach. The study of international relations may not be exact, given limitations noted by Wittgenstein and others, but precision is a practical research problem, not an insurmountable barrier to progress. In fact, most observers who point to the context-dependent nature of language are critical not so much of the social sciences but of the incorrect application of scientific techniques to derive overly precise measurement of weakly developed concepts. Clearly, our understanding of the causes of international conflict – and most notably war – has improved considerably as a consequence of applying sound scientific methods and valid operationalizations (Vasquez 1987, 1993). The alternative approach, implicit in much of the postmodern literature, is to fully accept the inadequacy of positivism, throw one's hands up in failure, given the complexity of the subject, and repudiate the entire enterprise. The most relevant question is whether we would know more or less about international relations if we pursued that strategy.

Cultural Relativism. Others have attacked positivism on the basis of cultural relativism, basically using an extension of the language arguments raised earlier. These critics reject any attempt “to secure an independent foundation, or Archimedean Point, from which to orient and judge social action” (George and Campbell 1990, 270). They consider it unfortunate that a majority of scholars, particularly in North America, “have not been made mindful of the way in which the affluence of [their] nation, as well as its size, has affected [the] research agenda” (Alger 1976, 70). While American academics are preoccupied with questions of power politics and crisis management,

Latin American and African scholars are concerned with dependence and social justice. Consequently, the findings of "affluent white citizens" are likely to be considered invalid "by the vast majority of the world who do not share these characteristics" (70).¹⁶ For postmodernists and scholars like Alger who sympathize with some of their arguments, then, pretensions of creating a universal science must be rejected.¹⁷

The study of international conflict in the United States and the Soviet Union after 1945 is often used to illustrate the potential impact of culture on theory. Whereas realism, as defined earlier, served as the prevailing paradigm in the United States throughout the postwar period, Marxist-Leninism was the dominant explanatory base in the USSR. American scholars traditionally accepted nation states as appropriate analytical units, while Soviet scholars viewed these entities as artificial manifestations of ongoing struggles between dominant and subordinate socioeconomic classes. By implication, proponents of Marxist-Leninism focused almost entirely on patterns of economic dominance within and among societies, as opposed to the behaviour of rational, unitary actors seeking to maximise foreign policy objectives defined largely in terms of security interests – the theoretical base of *realpolitik*. For realists, then, interstate conflict was, and continues to be, understood with reference to security dilemmas, crisis bargaining theory, power politics, deterrence failures, and so on. In contrast, Marxist-Leninism maintained that late nineteenth- and twentieth-century conflict – indeed all social behaviour – is a product of imperialism and forces central to the capitalist world economy. Since each academic community viewed the world through such different lenses, both developed distinct conceptual frameworks for explaining international affairs. Thus, claim the postmodernists, it is inherently impossible to develop the essential Archimedean point, or any readily transferable (that is, culturally neutral) foundation for knowledge construction, presumably a prerequisite for scientific progress.

Contrary to these assertions, however, there is ample evidence that Soviet and American scholars (and others, for that matter) had a great deal more in common than postmodernists acknowledge.¹⁸ By the mid-1950s the IR discipline in the Soviet Union had begun to modify its strict Marxist-Leninist social-class interpretation of international activity, and it embarked on a more North American research agenda, both in terms of the discipline's guiding assumptions (that is, a movement towards *realpolitik*) and methods (that is, a more rigorous scientific approach to explanation) (Lynch 1989).¹⁹ Although international conflict remained the core of Soviet analysis, and although imperialism continued to be the "prime cause" of such conflict, prom-

inent Soviet scholars now were claiming that “these actions are refracted on many levels and pass through many stages ... It is only by studying the ultimate and immediate causes responsible for international conflict, its nature and character, the forces opposing it, the circumstances which threaten to aggravate it, and the methods of possible regulation or prevention of conflicts that an analysis of international conflict can be made” (Yermolenko 1967, 47, 53).

In the final report of a roundtable organized in 1969 by the Institute of the World Economy and International Relations, a gathering of prominent Soviet scholars aimed at discussing the future of IR theory, several of the tasks facing the discipline, noted in the conclusion of their report, resembled the guiding principles of IR research in the United States (Lynch 1989):

- 1 increased study of the subject, method, and basic categories of IR theory;
- 2 a comprehensive focus on the system and structure; and
- 3 development of more sophisticated research methods based on mathematical modelling and improved methods for forecasting international events.

This self-conscious methodological attitude developed because the Soviets acquired a more significant international role, which produced an environment that was “inherently less threatening and, indeed, often ripe with opportunities,” but vastly more complex, “requiring a higher level of analytical expertise than before.” This reinforced the movement toward empirical content and methodological sophistication. The changes, in other words, were a consequence of shifts in power and status, *not* culture.

The example of the United States and the USSR is an important, if not crucial, case because it is so often cited to illustrate the impact of culture on theory and the limitations of positivism. Ironically, it serves equally well to disprove the point; notwithstanding variations in culture, ideology, political system, religion, history, and so on, there developed an apparent uniformity in approach, a standard conceptual and methodological framework that guided theory. More importantly, given the widespread acceptance of, and convergence towards, the liberal-democratic ideal throughout the world in the early 1990s, the behaviour of states and peoples are likely to become more comparable in the future (Fukuyama 1989) and, by implication, much more susceptible to scientific study. Cultural variations are not sufficient, therefore, to render meaningless an approach that downplays those differences (or assumes similarity) for purposes of theory.²⁰

Historical Limitations of IR Theory. The argument that scholars have failed to appreciate the importance of change and its impact on the universal applicability of social theories represents a third strand of the postmodern critique. In this case, the problem concerns the "conceit of scholars," the belief that their theories transcend history and that "what they know is as old as the world" (Cox 1986, 212–14). For postmodernists, the time-specific nature of social theory, particularly IR theory, has rendered cumulation in the field impossible; everything we know as theory can be understood only with reference to a particular historical period.²¹

However, a fundamental contradiction exists in the argument that the process of change is sufficient to render IR theory invalid and cumulation impossible. After all, the same process of change that presumably has rendered realism obsolete can, at some future point, create conditions that reestablish the paradigm's validity as a close approximation of reality and, by implication, reestablish its significance as a theoretical framework. The point here is that theories may not be universally relevant across time, but the universal applicability of a validated theory does remain constant.²² Ironically, postmodernists are guilty of the same crime for which they condemn realists: their interpretation of change is far too restrictive.²³

Furthermore, although the international system may have undergone significant change recently, it is not clear how these transformations are damaging to theory. Postmodernists must explain the impact change has on particular theories and go beyond arguments that change, in and of itself, is enough to invalidate an entire theory-building program. The process, after all, is not boundless: some things change more than others, and the impact of relatively stable forces is rarely considered. The existence and proliferation of nuclear weapons, for instance, is a systemic condition that is likely to remain fundamentally unchanged for some time. Research on the behaviour of nuclear powers in the past, as well as ongoing investigation of the behaviour of nuclear states in the future, will continue to be important topics.

CONCLUSION

Scholars agree that progress in the social sciences cannot be measured by the same standards as those applied to the natural sciences. While one can easily list important advances in gene splicing, artificial intelligence, space flight, heart surgery, robotics, fibre optics, and so on, one would be hard pressed to cite similar examples in, say, sociology or international relations. On the other hand, social scientists have a much harder time acknowledging progress when it has occurred. As

noted in chapter 1, this is especially true in IR, where scholars have yet to agree on methods of collecting and assimilating important contributions in the field.

The present study has attempted to address this problem in the context of superpower crisis management, in both the nuclear and conventional realms, by making cumulation a central theme of the project. The literature on deterrence and crisis management underscored problems with previous work on the subject and suggested new directions for aggregate testing. These alternative approaches revealed the diversity of game-theoretic propositions and other propositions generally overlooked by critics of rational choice and proponents of the dominant testing strategy, and they gave additional empirical content to nuclear and conventional deterrence theory.

Another important objective of the project has been to offer a few speculative thoughts on the possibility for conflict and crises following the Cold War. Keeping Mearsheimer's (1990a, 9) advice in mind, namely, that "those who venture to predict ... should proceed with humility, take care not to claim unwarranted confidence, and admit that later hindsight will undoubtedly reveal surprises and mistakes," I would argue that there is little reason to be optimistic. The system is becoming less secure precisely because we no longer have the stabilizing forces that were created by the ideological divisions of the Cold War. Of all the policy alternatives becoming available as a consequence of these global changes, the logic of nuclear (MAD) and conventional deterrence should remain as important considerations for those interested in establishing stable nuclear rivalries and controlling ethnic conflicts as we approach the twenty-first century. It would be dangerous, to say the least, to ignore the diagnostic and prescriptive qualities of deterrence and compellence theory if we are interested in managing nuclear rivalry or preventing the next ethnic war.

This page intentionally left blank

Notes

INTRODUCTION

- 1 The title is a variation of Mearsheimer's article "Back to the Future" in *International Security* (1990a). Mearsheimer proposed reviewing theories as a basis for predicting the stability of post-Cold War Europe. However, he dealt exclusively with hypotheses related to the changing nature of polarity in Europe, while the present investigation goes well beyond this by developing a more comprehensive approach to the cumulation of crisis-management theory.
- 2 Kuhn (1970) questioned the true scope of cumulation in the natural sciences as well, an issue to be discussed in more detail later in chapter 1.
- 3 The author's contribution to collaborative work has been used in chapters 3 and 5 of this project (Harvey and James 1992; James and Harvey 1992).

CHAPTER ONE

- 1 Some observers have pointed to the expansion of the field as yet another explanation for the lack of integrative cumulation in IR. Writing almost thirty years ago, Platig observed that developments in IR "may have outrun the capability of any one person to evaluate [progress] ... in adequate depth" (1966, 182). If this was not an accurate assessment of the complexity of the discipline then, it certainly is true today. In addition to the diversity of professional tasks, intellectual approaches, and research objectives (for example, historical, theoretical, methodological,

and policy-oriented approaches), the field today encompasses a variety of substantive foci, one of which will receive more attention in depth here: international law, international organizations and regimes, international political economy, international communication, comparative foreign policy, conflict, crisis, and war resolution, crisis management, crisis prevention, arms control and disarmament, security studies, and so on. Cumulation in the midst of this sectoral diversity, even if it is occurring, has become increasingly difficult to identify. Aside from a few very competent literature reviews and research surveys (Brecher 1977, 1993, chaps. 1, 3, 4; Tanter 1979; Zinnes 1980a,b; Hopple and Rossa 1981; Singer 1981; Haas 1986; Levy 1988, 1989a,b; Kal Holsti 1985), the community has never developed a method of compiling and evaluating theoretical accomplishments. Occasionally there is some indication (or measure) of progress through an assortment of “state of the art” compilations, but more often than not these are stocktaking, inventory projects that fail to adequately assess research in terms of its contribution to knowledge. Given that almost all literature reviews continue to have very little cumulative impact, a more systematic approach might help. The approach used here is to evaluate progress on the basis of a specific sector of the discipline, thus allowing assessment of a more manageable portion of the research program. The present study, for example, applies this strategy to crisis-management theory, as related to U.S.-Soviet rivalry. The objective is to compile, evaluate, and synthesize what we have discovered about the conduct of these states during intense crises in order to develop alternative testing strategies that build on previous work. The focus on cumulation is explicit, thus rendering any predictions about the impact of change on global nuclear stability more convincing.

- 2 Efforts to develop linkages across levels of analysis have been few and far between (Rosenau 1966; Singer 1981; Brecher 1993).
- 3 See Bueno de Mesquita (1980a) for more on this issue.
- 4 It is encouraging to note that the latest stage in the debate has focused on how differences over deterrence as a theory can affect selection of evidence (Lebow and Stein 1990; Huth and Russett 1990).
- 5 Indeed, among leading criticisms of research efforts over the last forty years, the following top the list: too many cases were used and were not subject to in-depth examination; not enough cases were used to permit valid generalizations; the study lacked the rigour of deductive theorizing and quantitative methodology; mathematical modelling and game theory are too abstract and require an empirical referent; the time frame for the analysis was inappropriate – either too long or too short; the statistical procedures were flawed, given the variables used; operationalization of the variables was not reliable or valid; the level of analysis was inappropriate; the level of analysis was appropriate but the inferences

were not, and so on. The assumption underlying such claims is that cumulation is a product of the techniques and methods employed and, by overcoming these technical obstacles, theoretical progress should follow. Unfortunately, most of these criticisms serve only to perpetuate counterproductive debates that contribute very little to the theoretical growth of a discipline.

- 6 Paige's (1968) application of the Snyder, Bruck, and Sapin (1962) decision framework has been cited as one of the most comprehensive investigations of American decision making during the Korean invasion of 1950. Using their framework, Paige analyzed organizational, informational, and normative aspects of the U.S. foreign policy process over a six-day period, confirming several propositions about decision making derived from political psychology. Two comprehensive case studies of decision making in World War I in 1914 and in the Cuban missile crisis of 1962 (Holsti 1965; Holsti, North, and Brody 1968) and several studies of crises in the Middle East (Jervis, Lebow, and Stein 1985) also contributed to the growing evidence against the rational actor model. Additional evidence from Janis' well-known work (1982) on groupthink demonstrated that leaders often censor information and alternative viewpoints in order to preserve the emotional comfort of group unity. These pressures for group conformity are particularly strong during international crises when the use of military force is an option.
- 7 See Fleming (1991) for a detailed discussion of conflicting evidence and other problems with earlier experiments.
- 8 See Brecher (1973, 1975, 1993) and Brecher and Geist (1980) on Israel 1967, 1973; Shlaim (1983) on the Berlin blockade; Dowty (1984) on the Middle East 1958, 1967, 1973; Dawisha (1984) on Prague Spring; and Jukes (1985) on Stalingrad.
- 9 In their work on communication and bargaining in conflict Snyder and Diesing found that, despite initial misperceptions, decisions in crises were usually the product of rational processes. The choice, in other words, was very close to what one would expect from a rational, expected-utility maximizer.
- 10 Some claim that these issues are inconsequential, because rational choice was never intended to serve as a theory of behaviour. It is an assumption that guides the theory-building process. Consequently, a theory should be evaluated on the logical consistency and empirical validity of propositions derived from it, not on the descriptive validity of the assumptions (Bueno de Mesquita 1985).
- 11 For a detailed evaluation of the Lebow-Stein research program, see the special edition of *Journal of Social Issues* (1987, 43(4): 73–154; especially Tetlock 1987; Fischeff 1987; Kolodziej 1987; Russett 1987; and Krell 1987).

- 12 The following studies were included among those reviewed by Lebow and Stein: Russett (1963), George and Smoke (1974), Organski and Kugler (1980), Huth and Russett (1984, 1988), and Kugler (1984). Their criticisms, however, can be applied to a number of other studies based on rationality models (e.g., Fink 1965; Bueno de Mesquita and Riker 1982; Weede 1981, 1983; Peterson 1986; Betts 1987). The central problems associated with implementing deterrence as a strategy were also discussed at length in George and Smoke (1974), Jervis (1979), and Jervis, Lebow, and Stein (1985).
- 13 This approach is not unlike meta-analysis (MA), a popular method in the social sciences (especially in psychology and sociology) designed to assess the progress of a research program. The difference is that MA measures the overall direction of the relationship between an independent and dependent variable, based on a composite of the significance tests generated by studies on the same phenomenon. There are several problems with using MA in the present study. For instance, the independent, dependent, and intervening variables across studies are often operationalized using a variety of indicators. These differences are seldom considered when creating the sample; poorly designed studies are thrown in with much better ones, as long as they deal with the same two variables. Also, and most unfortunately, the probability of publication is a partial function of the statistical significance of the results, which biases the mean-effect size in favour of finding a relationship.

CHAPTER TWO

- 1 My contribution to collaborative work has been used as the basis for this assessment of aggregate evaluation of nuclear deterrence (Harvey and James 1992).
- 2 Smith's taxonomy has not been adopted without modification. The works included by Smith that focus primarily on propositions about arms race research are not included in Table 2.1. Since the primary concern here is with the role of *nuclear* weapons in the deterrence process, Smith's use of time as a dimension will be excluded; most of the studies included in the table – and all of the relevant cases from the studies appearing in the lower right-hand of the third wave – focus on the post-1945 period.
- 3 Based on seventeen cases of deterrence between 1935 and 1961, Fink's analysis reinterpreted that of Russett (1963) and thus will not be reviewed here. Huth (1990) deleted two cases from the list used in Huth (1988a). Finally, Huth, Gelpi, and Bennett (1993) focused on escalation of great power disputes to assess the relative explanatory power of rational deterrence theory and structural realism. Because

they did not deal with testing propositions derived from nuclear deterrence theory, their study was excluded from the present analysis.

- 4 Classic examples include Wohlstetter (1962), Russett (1967) and Paige (1968).
- 5 Most of these models are based on concepts developed within the confines of game theory and emphasize the choice and utility structures of decision makers during deterrence situations (Kaplan 1958; Ellsberg 1961; Rapoport 1960; Snyder 1971; Doran 1973; Brams 1975, 1985; Bueno de Mesquita 1980a, 1981; Gauthier 1984; Powell 1990). Other models focus on the conditions of nuclear equilibrium and examine the stabilizing/destabilizing effects of various weapons systems (Snyder 1971; Kupperman and Smith 1971; Gillespie and Zinnes 1975; Smith 1981; Cioffi-Revilla 1983). The principal assumption underlying most of the above-noted models is that decision makers are rational expected-utility maximizers. However, very few investigations, with Bueno de Mesquita (1980a, 1981) as notable exceptions, provide empirical evidence to support that claim.
- 6 Other third wave, multistate studies of deterrence are excluded from Table 2.1 because they do not deal specifically with the question of *nuclear* deterrence. Some of the more prominent examples include Wright (1965); Narrol, Bullough, and Narrol (1974); Bueno de Mesquita (1980a, 1981); Peterson (1986); Smith (1981); Jervis, Lebow, and Stein (1985); and Lebow and Stein (1987, 1989a,b, 1990).
- 7 Examples of each type of failure appear in order of presentation: (1) the North Korean attack, 1950; the Chinese Communist invasion of Tachens, 1955; the Soviet intervention in Hungary, 1956; and the Cuban missile deployment, phase 1, 1962; (2) the Berlin blockade, phase 1, 1948; Taiwan Straits, phase 1, 1954–55; Quemoy, phase 1, 1958; and the Berlin Wall, 1961; (3) the Berlin blockade, phase 2, 1948; Taiwan Straits, phase 2, 1954–55; Quemoy, phase 2, 1958; Berlin, 1958; Berlin 1961; and the Cuban missile deployment, phase 2, 1962.
- 8 According to Craig and George (1990, 121), “Eisenhower stood aloof, despite all the talk from his administration of liberation of Eastern Europe, during the Hungarian revolution of 1956 when the Nagy government took itself out of the Warsaw Pact and called for help from the West. Far from attempting to deter Soviet military intervention in Hungary ... Eisenhower told Dulles to find a way of assuring Khrushchev that while the United States did not approve of the Soviet intervention, it would not interfere.”
- 9 They selected cases on the basis of the following conditions: “(1) the officials of state A (attacker) are seriously considering an attack on state B (protégé), which is allied with or deemed important by state C (defender); (2) key officials in the defender state must realize this;

(3) recognizing that an attack is a distinct possibility, the officials of the defender state explicitly or by movement of military force threaten the use of a retaliatory force in an effort to prevent the attack; and (4) the threats of the attacker and the defender states must be overt and clearly entail the use of military force.”

- 10 In contrast to the impact of economic linkages, the existence of a formal military alliance between defender and protégé “played no positive role, and, if not backed up by more tangible ties, actually worked against the success of deterrence” (Huth and Russett 1984, 524).
- 11 According to Fink (1965, 57), “Effectiveness is the behavioural outcome of the defender’s threat, defined in terms of the attacker’s subsequent action toward the protector – attack or no attack. Credibility refers to the attacker’s cognitive reaction to the defender’s threat – belief or disbelief. Since these two variables are not identical, there is no strictly logical justification for assuming a one-to-one correspondence between them.”
- 12 Huth and Russett (1988, 29) summarized the results of their analysis as follows: “deterrence is likely to succeed when the immediate or short-term balance of forces favors the defender, when any previous crisis involving the same adversaries resulted in stalemate rather than a clear victory for either, and when the military and diplomatic bargaining process is characterized by tit-for-tat or firm-but-flexible strategies rather than bullying or appeasement. The long-term balance of forces and the defender’s possession of nuclear weapons make little difference.”
- 13 Using an expected utility model, Bueno de Mesquita and Riker also found that “as the number of nations with nuclear weapons increases, the chance of bilateral conflict becoming nuclear initially increases, and then decreases to zero when all nations are nuclear armed” (1982, 288).
- 14 Cimbala’s (1988, 137) discussion of the (in)consistency between a state’s (1) declared policy, (2) operational or employment policy, (3) force development or acquisition policy, and (4) arms control policy should also be considered when assessing the operational criteria for a success or failure. Coding usually is based on declaratory policy (DP). Although DP often is a valid indication of actual intent, particularly in the midst of a crisis, it may be sufficiently different from operational or employment policy to render any judgments regarding success and failure problematic. Such inconsistencies may affect the validity of empirical assessments of deterrence theory and the role of nuclear weapons in the bargaining process.
- 15 If the initial retaliatory threat by state A (defender) triggers a crisis for state B (initiator), then state B may come to view itself as the legitimate defender of the status quo, provoking a counter-retaliatory threat against state A.

- 16 In defence of their data set Huth and Russett argue that a sample of cases “must be sufficiently alike in theoretical terms that they can be validly compared.” As they explain, “The sanction threatened by state A may be based on military or non-military means (economic or diplomatic, for example) and the action which state B may take in pursuit of its policy goals may entail the use of military force or be non-military in nature ... The fundamental differences in the policy instruments used by state A to deter and the types of policy actions being considered by state B suggest that theoretical propositions on success or failure in one class of cases cannot readily be applied to other cases ... State B may be vulnerable to military retaliation by state A but may not be vulnerable economically” (1990, 473). Craig and George provide a practical example of this complexity in their study of the Yom Kippur War of 1973. Despite diplomatic efforts, “the United States was unable to guarantee Israel’s security from a variety of Arab threats ranging from economic blockade to terrorist attacks to full-scale offensives by conventional forces ... [T]he deterring powers found themselves forced to reevaluate their commitments as the dissatisfied powers altered their strategies in accordance with perceived weaknesses in *these commitments*” (1990, 139; emphasis added).
- 17 Cimbala’s (1988, 137) discussion of the inconsistency between a state’s declaratory, employment, force acquisition, and arms control policies serves well to underscore many of these problems. Coding usually is based on declaratory policy. Although the latter is often a valid indication of actual intent, particularly in the midst of a crisis, it may be sufficiently distinct from operational or employment policy to render any judgments regarding success and failure problematic. This fact has important implications for empirical assessments of deterrence theory more generally.
- 18 Nasser was reported to have predicted the likelihood of an Israeli attack at 80 to 100 percent. However, the estimate was made just prior to Israel’s retaliation, when the probability of invasion was high. It says nothing of Israel’s failure, earlier in the crisis, to deter Nasser by convincing him that an attack would be imminent if Egypt began mobilizing its troops, which it did. In fact, given the prime minister’s public statements rejecting any military retaliation, Israel’s weak response to Nasser’s decision to block the straits of Tiran, and the pressure Israel was experiencing from the United States to avoid a war, Nasser could not help but question Israel’s resolve and, by implication, the credibility of Israel’s retaliatory threat. The intention here is not to defend one or the other position or to encourage endless debate over interpretation of the historical record but to establish that these debates exist.
- 19 Quoted in Orme (1987, 120).

- 20 For a thorough description of the relevant political, religious, and military factions that played a role in the Lebanese civil war, see Khalidi (1979), and Deeb (1980).
- 21 For a detailed account of events surrounding the decision to set up the military cabinet, see Rabinovich (1984, 44).
- 22 The evidence cited by Dawisha in support of an increasing perception of threat by Syrian decision makers was based on interviews with Abdulla al-Khani (Syrian deputy foreign minister), Adib al-Dawoodi (President's adviser on foreign affairs), and Ahmad Iskander (Syrian minister of information) between January 7 and 13, 1978. Dawisha (1978, 247) cited statements by President Asad on 2 December (1975) to support his assertions: "Arab unity as far as we are concerned takes priority over any other aim. It is an aim for which we are struggling continuously. The fact that we have not achieved much in this field so far does not mean that we feel pessimistic or have despaired. The issue, as far as we are concerned, is a permanent aim which requires continuous effort on our part." For a thorough account of the speech, see *British Broadcasting Corporation* (1975).
- 23 Asad's comments on 20 July 1976 provided a clear indication that he expected Israel to intervene, not only because of its security interests in Lebanon but because it was in Israel's larger political interest in the Middle East to intervene. Asad acknowledged that "A decisive military action (by Kamal Junblatt) ... would open doors to every foreign intervention, particularly Israel's intervention. Let us all visualize the magnitude of the tragedy which might ensue if Israel were to intervene and save some Arabs (Christians) from other Arabs (Muslims) ... The partitioning of Lebanon is an old Zionist aim" (quoted in Ma'oz 1988, 128).
- 24 Realizing that they were at a disadvantage militarily, the Christian Maronites were generally supportive of the Syrian reforms, whereas the Muslim-PLO (anti-status quo) coalition were convinced that the proposals would not produce an equitable distribution of political power in Lebanon.
- 25 Prospects for political gains were enhanced by Syria's diplomatic skills prior to the intervention. By the time Syrian troops invaded Lebanon, Syria had managed to rally wide public support. Even traditional Muslim leaders – both inside and outside Lebanon – favoured Syrian intervention to stop the fighting (Deeb 1980, 131). Internationally, France and the United States supported the intervention as long as the Syrian troops were used to protect Christian Maronites from annihilation. The loss of Egypt from its sphere of influence compelled Soviet leaders to guard against losing Syria as well. (Ironically, Asad believed that intervention would bolster his independence from the Soviet Union and improve U.S.-Syrian relations). Opposition to the interven-

tion was quite weak. Although Iraq moved forces to its eastern border, the military threat to Syria was insignificant. As for Egypt, the signing of the Sinai II agreement compromised its authority over other Arab states and the PLO in Lebanon, so most of its activities were diplomatic in nature.

- 26 Syria experienced several strategic foreign and domestic pressures prior to the invasion that might explain why Israel's best deterrent efforts were challenged. Specifically, the invasion was essential to prevent partition, establishment of a radical regime in Lebanon that might pull Syria into an unwanted war with Israel, and intervention and occupation by Israel of southern Lebanon. In addition, the timing was perfect. Prior to 1974 the distribution of capabilities in the Middle East was bipolar, with Egypt being the dominant Arab power. Following the Yom Kippur War, however, important changes took place in both the bases and distribution of power in the region. This diffusion of power "reflected a decline in the material and political strength of Egypt," largely a consequence of its strained relations with the Soviet Union, its principal arms supplier (Noble 1983, 56). As capabilities of other Arab countries began to rise, Syria acquired both the incentive and capabilities to carry out a more assertive foreign policy. Syrian leaders were experiencing a number of domestic pressures at the time. Important economic changes were taking place in Syria during the mid-seventies that shifted Syria's industrial capacity towards the oil industry (Lawson 1984). This shift had a detrimental impact on regions in the northern part of the country, the centre of Syria's cotton industry. As unemployment in these areas increased, so did opposition to the Ba'ath regime. Moreover, food shortages and a 30 percent inflation rate made capital increasingly scarce within the Syrian economy. Because the regime was unable to pump enough money into northern regions to appease the growing opposition, Syria's armed forces were required to suppress the violence, provoking further opposition and culminating in a wave of strikes, bombings, and assassinations. A military move into Lebanon, intended to impose a solution to the fighting there, became increasingly attractive to the Syrian regime. Such an operation could provide Syria's rulers with significant additional resources that they could use to their own domestic political advantage (Lawson 1984, 474). The most important of these resources were the capital held by Lebanese financial institutions, consumer products imported from Lebanon, and the facilities at the port of Beirut.
- 27 The overwhelming consensus is that, throughout the crisis, Syrian and Israeli decision making was both cautious and calculated. Each move was discussed thoroughly and efforts were made to clarify intentions to ensure that the crisis did not escalate out of control (Dawisha 1978;

Haley and Snider 1979; Evron 1987; Rabinovitch 1984). Syria's diplomatic skills were effectively utilized to reassure both Israel and the United States that their interests would not be jeopardized. This was no easy task; strategies of reassurance are rarely successful in a setting of protracted conflict because conciliatory moves tend to be viewed with great suspicion. Notwithstanding Syria's diplomatic skills, however, Syrian leaders were mistaken in their judgment that the problems in Lebanon could be solved by a resort to force. Although a more massive intervention (three months later) achieved both a cease-fire and a political compromise at the Riyadh conference (October 1976), neither lasted very long.

CHAPTER THREE

- 1 The author's contribution to collaborative work has been used to describe coding procedures and selection criteria in this chapter (James and Harvey 1989, 1992).
- 2 It should be noted that nuclear deterrence theory (NDT) refers to a set of logically related axioms (based on rational choice) that combine to create a theory about nuclear war avoidance between two powers faced with the prospect of mutual annihilation. A number of studies, including many reviewed in chapter 3, confuse tests of NDT with tests of the *utility of nuclear weapons as a source of power and influence in international politics*. Although such tests are not sufficient – in and of themselves – to evaluate the theory, they do provide an appropriate starting point. After all, NDT assumes that the terror created by the threat of devastation is the key to preservation of peace (Kugler 1984).
- 3 Peterson (1986), while arguing against the notion that it is intrinsically easier to deter than to compel, pointed out that the practical difference between the two may be exaggerated anyway.
- 4 Cases in progress by 1949 may have started earlier; the first two cases in the table actually began in 1948.
- 5 The five indicators of severity of threat, described in chapter 5, were selected to tap into forces that are expected to play a major role in the decision calculus, especially when estimating the consequences of crisis escalation.
- 6 Of course, additional evidence with respect to the first two propositions can be obtained by extending the data set to include other nuclear rivalries (for example, USSR-China 1969) or nuclear/non-nuclear and non-nuclear/non-nuclear cases. However, the emphasis on the longest nuclear rivalry in history represents the core of the present investigation. A more comprehensive test using additional cases is planned, but is beyond the scope of this study. In any event, the addition of a few

outlier cases (for example, USSR-China 1969 or China-USSR 1979) is not expected to significantly alter the overall findings or conclusions reported here.

CHAPTER FOUR

- 1 The author's contribution to collaborative work has been used in analyzing the four game-theoretic models of strategic choice (James and Harvey 1992).
- 2 The matrix below represents the game in its basic form. Two drivers, Row and Column, approach each other on a road. The objective of the game is to force the adversary to swerve out of the way. The payoffs in a given cell correspond to those of Row and Column, respectively. Relative magnitudes are $r_4 > r_3 > r_2 > r_1$ and $c_4 > c_3 > c_2 > c_1$. If each swerves, both receive the second-best payoff. If Row/Column swerves and Column/Row does not, Column/Row obtains the best payoff and Row/Column the third best. When neither swerves, both receive the worst payoff.

The Game of Chicken

		Column	
		Swerve	Do not swerve
Row	Swerve	r_3, c_3	r_2, c_4
	Do not swerve	r_4, c_2	r_1, c_1

- 3 Regardless of the distinction between conventional and nuclear deterrence, however, some have claimed that PD is still the more appropriate analogy. As Zagare (1985, 160) observed, "whatever the explanation for the popularity of the Chicken analogy ... it is clear that it tells only part of the story. Chicken may very well encapsulate the problem of mutual deterrence, especially in the nuclear age, wherein deterrence threats are inherently incredible. But if, as the theory of deterrence suggests, stable mutual deterrence requires that both players possess a credible threat, then Prisoner's Dilemma represents the *solution* to the problem of mutual deterrence." By incorporating the notion of *nonmyopic equilibrium* (Brams and Wittman 1981) and altering the assumptions associated with the payoff structures, Zagare concluded that PD was the more accurate model. However, using the nonmyopic equilibrium instead of Nash may simply have the effect of transforming what seems to be PD into chicken. The question then becomes which equilibrium is more realistic.

- 4 As Brams and Kilgour (1988, 103) point out, “Axelrod’s conclusions concerning the robustness of tit-for-tat ... are in large part empirical and follow from a model that assumes repeated play against different opponents. By comparison, we propose an analytical model in which aggressive actions are immediately known to the players, there may be varying levels of aggression and retaliation, and there are two rather than many different players.”
- 5 The substantive domain of their investigation included two studies of 200 subjects each. The evidence was based on tests in which “subjects played trials of Prisoner’s Dilemma with a simulated other who could send verbal communications ... The other began intensifying conflict by being behaviourally unresponsive over a long duration, by being totally noncooperative, by responding more quickly to the competition than the cooperation of the subject, or by coercing cooperation with threats for the purpose of exploitation. Then the other introduced a program of carefully communicated conciliation” (1986, 99).
- 6 From his research on protracted social conflict, for example, Azar (1986, 31–2) concluded that “conflictual and cooperative events flow together even in the most severe of intense conflicts ... However, conflictual events are clearly more absorbing and have more impact on determining the consequent actions of groups and nations. Cooperative events are not sufficient to abate protracted social conflicts. Tension reduction measures may make the conflict more bearable in the short term, but conflict resolution involves a far more complex process than mere conflict management.” Azar also pointed out that in some cases, traditional processes of tension reduction (for example, power bargaining and mediation) may themselves be a reason for conflicts to be protracted; they “lead to temporary settlements without tackling the underlying issues” (52).
- 7 Indeed, implicit (and sometimes explicit) in most game-theoretic research, including the rationality models developed by Axelrod and Brams and Kilgour, is the assumption that a transition from conflict to cooperation will be problematic.
- 8 Lindskold, Betz, and Walters (1986, 100) cite the example of the trip by President Sadat of Egypt to Jerusalem in 1977, “a significant and surprising act to the Israelis who viewed him and Egypt as the enemy. In retrospect, the Egypt-Israeli peace treaty came along rather promptly” once this initiative was offered. Additional case studies in the realm of superpower relations also indicated that “a very quick response” usually follows the introduction of GRIT. Consider the following example: “The 1963 Kennedy Experiment began with the *Strategy for Peace* speech at the American University in June and was followed by initiatives on testing the hotline, wheat sales to Russia, opening consulates, and air

travel, and so on. Khrushchev replied almost immediately with the publication of Kennedy's speech, conciliatory actions, and a speech of his own" (101).

- 9 Goldstein and Freeman (1990) presented in-depth studies of six cooperative initiatives involving China, the USSR and the United States: Zhou's Bandung initiative (April 1975), Zhou's trip to Eastern Europe (January 1957), Johnson's gestures of goodwill (October 1966), Nixon in China (1972), Gorbachev's testing moratorium (July 1985), and Gorbachev's Vladivostock speech (July 1986). The intention was to illustrate, through examples, how sustained cooperation between major powers is enhanced through implementation of GRIT-type strategies. Consider, for example, the USSR: since 1985 it "has implemented a strategy that strongly resembles super-GRIT. Building on a history of growing Soviet reciprocity over the past twenty years, Soviet leaders have launched a series of (apparently ongoing) cooperative initiatives toward both the United States and China" (154). These actions include the moratorium on nuclear tests in July 1985, the Vladivostock concessions to China in 1986, a phased withdrawal from Afghanistan, Gorbachev's UN speech in December 1988 on unilateral military cuts and the unilateral withdrawal of 500 nuclear weapons from Europe.
- 10 Dacey and Pendegrift (1988) found that GRIM performed quite well in a computer simulation that included several other strategies.
- 11 One further point regarding Mueller's depiction of GRIM should be noted. Even if it is true (as Mueller argued) that GRIM comes out ahead when the opponent misinterprets a noncooperative act, the probability that noise would play a role in this regard depends on how cooperative (or noncooperative) the act is. It would be more difficult, in other words, to misinterpret a noncooperative act (because of noise) if that act was unambiguous (for example, a full-scale conventional war). Conversely, the probability that noise would cause the opponent to misinterpret a defection would be greater if the bargaining technique was relatively more ambiguous (for example, an economic, political, or diplomatic act); as the difference between provocation and retaliation increases, ambiguity decreases and so does the probability that noise will play a role.
- 12 The values along the axis *decline* as t_i and $q(t_i)$ become more intense, with the polar points being 1 (status quo, or nonpreemption) and 0 (extreme hostility, or maximum preemption).
- 13 Economic depression in the 1930s, for example, "stimulated demands for protection by firms and individuals in distress, and therefore reduced the incentives for governments to cooperate with one another" (Axelrod and Keohane 1985, 228). The payoff for mutual defection, in other words, became higher than the payoff for mutual cooperation,

thus altering the structure of the game. Axelrod and Keohane (228–9) cited Jervis' (1978) study of the shift from balance-of-power to concert as additional evidence. "After world wars, the payoff matrix for the victors may temporarily be one of 'stag hunt': fighting together results in a short-lived preference for staying together. After a war against a hegemonic power, the other great powers often perceive a mutual interest in continuing to work together in order to ensure that the defeated would-be hegemon does not rise again. As recovery from the war proceeds, one or both parties may come to value cooperation less and relative gains more. And if one side believes that its counterpart prefers to defect, its own preference will shift to defection in order to avoid the worst payoff [mutual defection]." Consider superpower crisis bargaining: given the environmental effects of *any* nuclear explosion, and the likely repercussions with respect to relations with other countries, the United States and the Soviet Union may prefer mutual cooperation to unilateral defection in any crisis at the brink of nuclear war, a change from chicken to stag hunt ($cc > c'c > c'c' > cc'$). Again, this preference ordering would largely depend on the specific actions taken by the opponent.

- 14 Another criticism is that game-theoretic models assume that actions are perceived and correctly acknowledged by the participants. But the effects of aggressive acts may be extremely difficult to assess, as can be seen in the example of damage reports. As Beer observed (1986, 183), "Actors are ignorant in varying degrees of others' behaviour. Part of the trouble comes from simple confusion ... Some of the difficulty has come from advances in technologies as well as the expansion of public and private bureaucracies ... Inconsistencies and contradictions occur. Much information is spurious, the result of sloppiness, concealment, and partial truths or lies." If this is an accurate depiction of states in the real world, then tailoring threatened punishment to levels of provocation or aggression will more than likely be difficult. In other words, it is doubtful, given these limitations, that actors in a dispute could decide either what the optimal (proportional/nonproportional) response to the particular aggressive act should be or whether that response would be perceived by the opponent as proportional, more than proportional, or less than proportional. Indeed, this criticism is applicable to all four models. Even if they are logically consistent, it still is unclear whether states are capable of acting in the manner prescribed. Achen and Snidal (1989, 164) offer the following response: "Rational deterrence theory is agnostic about the actual calculations decision-makers undertake. It holds that they will act as if they [have] solved certain mathematical problems, whether or not they [have] actually solve[d] them. Just as Steffi Graf plays tennis as if she did rapid computations in Newtonian

physics ... so rational deterrence theory predicts that decision makers will decide whether to go to war as if they did [the] expected utility calculations."

- 15 Of course, that effectively raises the question of which among the four models of response provides a more realistic account of behaviour in crisis situations. With respect to the Brams and Kilgour model, for example, the question is whether decision makers act as if they are capable of deciding what the optimal (proportional/nonproportional) response to a particular aggressive act should be and whether that response would be perceived by the opponent as proportional, more than proportional, or less than proportional. Similar questions apply to the other three models.
- 16 James (1991) used two payoff configurations in an aggregate analysis of retaliation to provocation in superpower crises. One version contained identical values for the United States and the USSR while the other preserved ordinal symmetry but altered the size of the intervals. James found that agreement with a threshold value for satisfactory response could be enhanced by assuming that the United States placed a higher relative value on the preplay position. However, his approach still incorporated only four values for c_2 and c_3 among the vast array that might have been selected, thus making it appropriate to explore the other options listed in Table 4.1.
- 17 For a more detailed description of coding procedures for provocation and retaliation levels, see Brecher, Wilkenfeld, et al. (1988).
- 18 Several of the cases from James and Harvey (1989) have not been included in the present list. James and Harvey focused on satisfactory, as opposed to optimal, responses to provocation in superpower crises. In that testing environment the level of direct involvement of the superpowers was not considered to be crucial, because the behaviour of client states (that is, bloc members) could be permitted to play a greater role. In order to assess the models of retaliation in terms of relative deviation from a precise level of response, cases in which one of the superpowers played a less prominent role have been excluded. The relevant crises are Pleiku, the invasion of Laos, the mining of Vietnam ports, the Christmas bombing of Vietnam, and the Soviet threat to Pakistan. In the first four cases the USSR is judged to play a secondary role, while that is regarded as true for the United States in the final crisis.
- 19 The ICB project presently is in the process of releasing data on cases up to and including 1988. It also should be noted that 1985 is the last year before the era of reform in the USSR.
- 20 Five of the scale points from the previous study have been eliminated: internal physical challenge to the regime; external change, political and

- military; internal verbal challenge to the regime; and others. These categories are not relevant, in empirical terms, to strategic choices in a direct confrontation of the superpowers in a crisis situation.
- 21 A more general form of this scale, applying to states below the superpower level, would have to allow for internal characteristics such as the level of development. Some scale points might be relatively more intense for a target with a developing economy or an unstable government. For example, in comparison to nonviolent military actions, economic and political acts might be experienced more intensely by such states. Even if the order of the scale points turned out to be invariant, the intervals along the way might be subject to change.
 - 22 Specific acts of provocation and retaliation sometimes will consist of observed behaviour as opposed to verbal or written assertions. In order to deal with this variation, a scale of coercion would have to grant more credibility, all other things being equal, to an action rather than a threat. More specifically, violence should be accorded more weight in coding decisions than nonviolent military actions, which in turn would stand above threatening statements. In sum, actions speak louder than words.
 - 23 It might be suggested that identifying the act of noncooperation that initiates a crisis requires problematic value judgments. However, the coding procedures of the ICB project, described in chapter 3, minimize problems of reliability. Also, with respect to triggering acts, it is obvious that neither cooperation nor a lack of action (the final two scale points) can trigger a crisis. However, superpower rivalry is assumed to be in progress at all times. Each of these nonthreatening forms of behaviour may thus be regarded as a choice available to the players at any given moment.
 - 24 It should be noted that the threat and retaliation moves are sometimes separated by more than just a few days, raising the possibility that other events could intervene and contaminate the process. However, use of the ICB project's major response variable to code $q(t_i)$ counteracts that possibility. The major response corresponds to the salient words and deeds directed specifically toward the party held responsible for triggering the crisis. Other actions by the retaliating power are not incorporated in the measurement of $q(t_i)$.
 - 25 The descriptions that follow are based on Brecher, Wilkenfeld, et al. (1988, 218–20, 293). Both "satisfaction" and "goal achievement" are treated as ordinal variables with the following categories. Satisfaction: 1 = defender satisfied, adversary dissatisfied; 2 = all sides satisfied; 3 = all sides dissatisfied; 4 = adversary satisfied, defender dissatisfied. Goal achievement: 1 = victory; 2 = stalemate; 3 = defeat. The decision to rank "adversary satisfied, defender dissatisfied" fourth is consistent

with expectations of Soviet-American preferences within the confines of the Cold War.

- 26 Kolkowicz (1986), for example, characterized the Soviet view of the geostrategic status quo as less favourable than that of the United States. While the ordering of payoffs would be preserved, the emphasis on stability – so prevalent in the West – would be much lower.
- 27 Among the corner payoffs, the utilities for all but (r_1, c_1) could be allowed to show variation from one crisis to the next. In testing the revised model, however, it will be assumed that the corner payoffs are constant across cases.
- 28 It is possible that every national leader will attach a very high value to the status quo, (c_3, r_3) , perhaps close to 1.0. However, the occurrence of international crises suggests that the USSR and the United States are sometimes willing to risk war in order to obtain concessions (although intending to return the situation to the status quo). Thus the lower values attached to the status quo by the USSR and the United States, $2/3$ and $3/4$ respectively, are viable choices based on past performance.
- 29 The tau b statistic measures the percentage reduction in error in predicting categories of the dependent variable. For example, if tau $b = .25$, then the independent variable in question produces a 25 percent reduction in error when predicting categories of the dependent variable. It should also be noted that with a population rather than a sample of cases significance levels appear in an advisory role.
- 30 Goldstein and Freeman also presented six case studies of cooperation between each dyad (for example, Zhou's trip to Eastern Europe in 1957, Nixon's visit to China in 1972, and Gorbachev's testing moratorium in 1985). The purpose was to furnish historical support for Osgood's (1965) claim that cooperation breeds cooperation, and set the stage for the aggregate analysis of alternative GRIT-type strategies in subsequent chapters.
- 31 By contrast, the approach used in this chapter measures effectiveness in terms of escalation control and subjective and objective assessments of outcomes – defined in terms of "satisfaction" and "victory/defeat" respectively. For reasons to be discussed below, this technique is expected to provide a more accurate assessment of the models.
- 32 The logic of PGRIT is that one side starts with a cooperative act (usually symbolic) and increases the level of cooperation in subsequent moves until a settlement is reached. The empirical investigation outlined in Goldstein and Freeman actually assesses PGRIT and GRIT simultaneously and does not differentiate between them during the stage of testing. In any case, it is not the particular variation on the model that is important but the effectiveness of cooperation in general as a crisis management technique.

- 33 A similar argument was presented by Ray (1991, 1494). “Each of the three countries analyzed has probably been more likely to be the target of cooperative acts during and shortly after times when it has acted in a cooperative manner than at other times.” The question Goldstein and Freeman failed to address was whether the cooperative initiatives caused cooperative responses or whether they were simply the result of changes in the relationship.
- 34 Chapter 1 of this study provides a detailed evaluation of evidence supporting this claim.
- 35 The same defence was used as partial support for Goldstein and Freeman’s findings on reciprocation; if reciprocation did not occur, one would expect to find no observable patterns. If such patterns develop in the relationship, then all the more reason to accept the reciprocity thesis.
- 36 See also Goldstein (1991).

CHAPTER FIVE

- 1 For a detailed summary of coding decisions for key variables, see Wilkenfeld, Brecher, et al. (1988).
- 2 The authors cite the 1956 Suez crisis as an example: “The United Kingdom and France were major powers in global terms, and Egypt was at most a middle power. The former countries were economically advanced; the latter had a very poor developing economy. The adversaries differed in political regime – Western democracy versus military authoritarian – and in culture – language, belief system, and so on. These differences were accentuated because the United Kingdom and France were struggling to preserve existing international system rules regarding rights embodied in interstate agreements over the Suez Canal, whereas Egypt was asserting national rights as a basis of a higher rule” (1986, 38). The decision to implement a military solution to the crisis was a product of these cleavages. In fact, the political regime and culture of Israel and Egypt created cleavages that remain central to many of the continuing problems of the region.
- 3 “Intensity” refers to the level of violence and hostility employed by the actors as a crisis-management technique. “Centrality,” on the other hand, refers to “the relative importance which the decision makers attached to their use of violence in order to attain their goals in the particular crisis” (Wilkenfeld, Brecher, et al. 1988).
- 4 The tau *b* statistic measures the percentage reduction in error in predicting categories of the dependent variable. A tau *b* of .25 means that the independent variable produces a 25 percent reduction in error when predicting categories of the dependent variable. It should also be noted

that, with a population rather than a sample of cases, significance levels appear in an advisory role.

- 5 Of the nine crises in which very high threats occurred, 78 percent (7) exhibited minimal levels of violence.
- 6 Included in ρ_2 is the assertion that nuclear powers should achieve draws when directly confronted with other nuclear powers. To test this, only cases of mutual deterrence were examined. Of these eleven crises, only four (36 percent) provided evidence to support that aspect of the proposition (that is, the dispute ended in compromise or stalemate). In other words, a majority (64 percent) of the cases with both superpowers directly involved as crisis actors ended in a defeat/victory scenario.
- 7 It should be noted that the term “nuclear deterrence theory” refers to a set of logically related axioms based on rational choice that combine to create a theory about the avoidance of nuclear war between two powers faced with the prospect of mutual annihilation. A number of studies confuse tests of nuclear deterrence theory with tests of the utility of nuclear weapons as a source of power and influence in international politics. Although such tests are not sufficient – in and of themselves – to evaluate the theory, they do provide an appropriate starting point.
- 8 This is consistent with the pseudocertainty effect noted by Kahneman and Tversky (1979). See also Farnham (1992, 226).

CHAPTER SIX

- 1 Although there are numerous historical case studies that claim to refute such findings (for example, Foot 1989), there is an equally impressive body of historical evidence that supports them (for example, Dingman 1988). In her analysis of Eisenhower’s nuclear threats in the Korean War, Foot (1989, 107–9) questioned the assertion that they were instrumental in bringing about an end to the conflict. According to Foot, domestic economic factors (for example, North Korea’s dismal economy and the beginning of China’s first major five year plan) greatly influenced decision makers in both countries to move toward compromise and peace talks. However, Foot conceded that these domestic forces can only be partially linked to the end of the conflict. Economics, in other words, cannot explain why the end of the Korean War came when it did. Dingman’s (1988) assertion that the threat of nuclear retaliation acted as the major catalyst to ending the war is consistent with the findings of this study.
- 2 For a comprehensive critique of the “long peace” literature, and an important refinement of the concept of global instability, see Brecher and Wilkenfeld (1991).

- 3 These results support Oneal's (1988, 605–9) contention that rational choice frequently will prevail in crises: "The centralization of authority, decline in partisan politics, and greater acceptance of the leader's responsibility that typically occur during crises create an opportunity for leadership and innovation." The concentration of authority also suggests that fewer independent interests require accommodation. Furthermore, senior officials are usually generalists, with less inclination to bargain in order to protect bureaucratic interests.
- 4 For a discussion of these unwritten rules see Zacher (1992, 72): "do not threaten the second strike capability of the other side (i.e., support mutual deterrence) and, in fact, seek to enhance it (e.g., ABM treaty, 1972); use nuclear weapons only as a last resort when one's territory or that of one's core allies [is] threatened; avoid direct military conflict with the forces of other great powers; do not militarily threaten the core allies of other great powers; do not undermine the ability of other great powers to monitor major military activities; prevent allies from undertaking actions (particularly against client states of the other side) that could drag one into a great power war," and so on. For arguments about how U.S.-Soviet nuclear rivalry "widened the gap between the value of the interests in conflict and the possible cost of war" and increased the "range of manoeuvring" that promoted crisis management and de-escalation, see Dougherty and Pfaltzgraff (1990). See also Brodie (1959a,b), Schelling (1966), and Waltz (1981) for similar arguments.
- 5 The same applies to other studies (for example, Lebow and Stein 1989a,b, 1990; Taylor and Ralston 1991) in which nuclear weapons or nuclear deterrence was not an issue. It should be noted that Brecher and Geist (1980) and Dowty (1984) have provided extensive evidence to support a rational decision-making model in Israel and Egypt during the 1967 and 1973 wars.
- 6 For a review of arguments by nuclear optimists and pessimists, see Sagan (1994) and Feaver (1994, 1993).
- 7 Feaver (1994) makes the same point in his review of arguments by optimists and pessimists.
- 8 Twenty-five tons of plutonium, produced for the Soviet breeder program, is no longer needed. "This plutonium is just sitting there, not being used for anything. The incentive for the central government to sell some of it clandestinely is very large" (Ellsberg 1992, 141).
- 9 See Molander and Wilson (1994) on the failure of the IAEA and other nuclear-intelligence collection capabilities. For a recent discussion of the inevitability of proliferation in the Middle East, see Kahana (1994).
- 10 Feaver (1993, 182) pointed out that these recommendations are not new and were actually considered by U.S. policymakers as early as the

1960s. See also Larus (1967) for an early set of policy proposals on nuclear safety and a “common defence” that recommend sharing command-and-control technology.

- 11 Sagan (1994) acknowledges two problems with this strategy: (1) it gives the impression to prospective nuclear clients that the United States is ready to support their proliferation efforts, and (2) officials within the new nuclear state may not be willing to give U.S. experts access to their installations.

CHAPTER SEVEN

- 1 Although this distinction is useful, there are exceptions that should be noted. If state A threatens retaliation to contain the spread of violence in a crisis, does this constitute an attempt to deter escalation or compel compliance with demands to keep the fighting to a minimum? Similarly, does one deter a state from rejecting the latest peace proposal or compel leaders to accept it? Did the United States threaten Serbia with economic sanctions in order to compel its leader, Milosevic, to endorse the Vance-Owen plan, deter undesired actions in the form of Serbia’s rejection of the accord, or compel Milosevic to deter Karadzic from escalating Bosnian Serb attacks on Muslim enclaves? These questions are especially relevant in the case of the former Yugoslavia, because the coercive threats attempted to accomplish more than one objective: deter escalation, particularly with respect to ethnic cleansing, through the creation of safe havens and no-fly zones, and compel the Bosnian Serbs to accept Vance-Owen or, at least, return to the bargaining table. Even a straightforward threat linked to protection of safe havens is problematic – were the Americans and Europeans (through NATO) trying to compel the Serbs to back away from territory bordering these safe areas or to deter them from crossing the exclusion zone and shelling the city, or both? The issue becomes even more confusing when one considers encounters that constitute examples of successful deterrence but failed compellence, and vice versa. A major advantage of looking at a crisis in terms of separate and distinct exchanges is that it offers an opportunity to look for both types of interaction in the same case, thus avoiding the need to make a judgment about motives and intentions across the entire crisis.
- 2 The approach may bias case selection towards identifying failures, but if behaviour in a crisis is consistent with expectations, notwithstanding this selection bias, all the more reason to accept the theory.
- 3 The crucial question is whether the retaliatory threat is more costly to the challenger than the sacrifice incurred through capitulation or whether the sacrifice is significant to the challenger’s overall objectives?

For example, the no-fly zone succeeded, but the Bosnian Serbs simply switched to a ground attack and began operating in Serbian air space. Since the objective was to deter the use of air space over Bosnia, the threat worked. The no-fly zone resolution was never expected to end the war or create peace in the region.

- 4 A few notes about commitment are in order. Rational deterrence theory stipulates that the challenger will assess the costs and benefits of inaction versus action depending on, among other things, a critical judgment by the challenger of the defender's commitment to its threat. However, cases cannot be selected on the basis of whether there was a serious commitment by the defender to retaliate, because this would naturally bias the results in favour of successful deterrence (see Fearon (1994) for an excellent discussion on this point). When assessing the defender's commitment to respond, a challenger will consider the severity of its challenge in terms of the defender's political and security interests. Consequently, challenges are often tailored in such a way that retaliation is deemed by the defender to be unnecessary, as demonstrated by the use of limited probes by the Bosnian Serbs to test NATO's resolve (see George and Smoke 1974).
- 5 This time frame involved the most diplomatic and military interaction between the two sides since the onset of hostilities; it provides a large body of evidence for evaluating the theories.
- 6 Using the *New York Times* and *Keesing's* ensured that both American and European perspectives were considered when recording interactions and events. Major inconsistencies in reports were noted. In general, *Keesing's* was more vague on details; whole exchanges (for example, exchange 3) were even excluded from coverage. Several additional American and European sources were used to confirm *New York Times* and *Keesing's* coverage – *U.S. News and World Report*, the *Economist*, the *Herald Tribune* (London) and the *Times* (London). Coverage in these journals tended to be less specific and more policy-oriented, focusing on the technicalities of the latest peace accord, while providing fewer detail on interactions.
- 7 Two senior policy analysts within NATO were consulted – Edgard Vandeputte (Defense Planning and Policy Division, Crisis Management Section) and Richard Zandee (Disarmament, Arms Control, and Cooperative Security Section, Political Affairs Division). Each analyst was asked to review the eight exchanges and to comment on both the accuracy of the historical account and the author's interpretation of perceptions and motives. Their recommendations and suggestions have been incorporated in the final presentation of the encounters. Major-General MacInnis was asked to review the exchanges as well. For many exchanges, the general was either directly involved or reviewed docu-

mentation and strategies with those who were. From his office in Zagreb he had direct contact with NATO officials in Brussels and Naples as well as with UN officials in Zagreb and New York. In many instances he served as one of the bridges between the two organizations. As a Canadian, he had access into both camps, unlike Under-Secretary General Akashi, who was unfamiliar with NATO's civilian structure or the force commanders (Cot and deLepresle), who were not familiar with the workings of the integrated military structures.

- 8 For a detailed review of the exchanges, along with summaries of events, descriptions of threats and responses, coding of the severity and credibility of the threat, and an evaluation of the outcome in terms of the theory's predictions, please see Harvey (1996a,b).
- 9 I am indebted to Gary Geortz and Bear F. Braumoelle for advice on how to apply the necessary and sufficient condition hypothesis to deterrence success. For an excellent discussion of the issue, see Geortz and Braumoelle (1996).
- 10 Evidence that $YYYN$ never occurs suggests that the absence of at least one condition is necessary for deterrence failure. If there are cases of $YYYY$ and no cases of $YYYN$, the evidence supports a claim of sufficiency. "If you have one condition which is necessary, you can not have another which is sufficient. If sunlight were necessary for plant growth, there would be no other condition which could be sufficient for plant growth, because that would mean that it could make plants grow in the absence of sunlight – which we know can't happen. It's like an unstoppable force and an immovable object: if one exists, the other can't. These properties extend to jointly necessary and jointly sufficient conditions as well – if you have a set $S = \{x_1 \dots x_n\}$ of conditions whose presence is jointly sufficient for success and a set $N = \{x_1 \dots x_m\}$ of conditions, one of whose absence is necessary for failure, it must be the case that everything outside the intersection of those sets can be safely dropped" (Bear F. Braumoelle, personal correspondence, 1996).
- 11 James (1993, 23) also noted that "there still is considerable impressionistic evidence in favour of state-centrism," citing Deutsch's (1988, 45) observation regarding the "25-1 (aggregate) ratio of state spending power as compared to that of international organizations."
- 12 Even proponents of complex interdependence accept this line of reasoning. In fact, the disagreement between paradigms often has less to do with the limitations of state-centrism as a unit of analysis (or with the assumption that anarchy, power politics, and conflict are defining characteristics of the system) and more to do with claims that realists accept these assumptions as inevitable and unchanging laws of nature. Ashley (1986) and Cox (1986) are two examples of critics who attempt to give realism a character and a set of assumptions that it does not necessarily

hold. The collective response by Waltz (1986, 339), Gilpin (1986, 316–17), and Keohane (1986) to Ashley (1986) on this point serves to illustrate the extent of misunderstanding implicit in such assertions and the counterproductive nature of these debates. As Waltz (1986, 339) concluded, “the distinction between an assumption and a statement striving for descriptive accuracy should be easy to grasp.” States may not really exist in concrete form, “but neither do Allison’s bureaucracies, interest groups, nor even transnational actors ... Only individuals exist [and] ... only individuals act” (Gilpin 1986, 318). Our theories proceed from the assumption that these social and political entities exist as a “matter of convenience and economy.” The alternative to using state-centrism and rationality (or any other simplifying assumption) as a basis for theoretical development is to give up and throw out the entire enterprise (Waltz 1986).

- 13 Some have claimed that a more accurate depiction of international politics could combine elements of both (Ruggie 1983; Keohane 1984, 1986b; Rosenau 1992). For instance, it may be appropriate to accept state-centrism as accurate when military-security issues dominate the agenda (that is, when security is scarce – during times of international conflict, crisis, war, deterrence, and so forth) but to acknowledge at least some softening of the “hard shell” with respect to other issues (for example, economic, cultural, scientific, and environmental issues – traditionally West-West issues). Of course, even in situations where the direct use of force has essentially been barred, in West-West conflict, for example, military power has still played an important role: the United States gained a great deal of bargaining advantage vis-à-vis Western Europe as a direct consequence of its deterrence strategy. In any event, it is essential to determine the applicability of realism or complex interdependence to each situation (Keohane and Nye 1977, 28–9). Without this determination, further analysis is likely to be incoherent.
- 14 The terms “poststructural” and “postpositive” have also been used to describe this intellectual movement. Postmodernism and poststructuralism refer to slightly different strands of critical theory. However, the purpose here is only to extract the broad criticism implicit in both schools. Rosenau (1992, 3) distinguishes between the two in terms of emphasis, arguing that “postmodernists are more oriented towards cultural critique, while the poststructuralist emphasizes method and epistemological matters.” George and Campbell (1990, 281) noted additional differences between hermeneutics and poststructuralism. While hermeneutics and other critical theories expose “the rotten foundations and the ideological function of traditional theory ... to enable people to overcome the power structures that oppress them,” poststructuralism focuses “less on attempting to secure emancipation through the unmask-

ing of power, oppression, and ideology, and more ... on concrete examples of the way power is used in all of society's sites ... and how to effectively resist the imposition of power articulated via the privileged 'logocentric' discourses of modern scientific rationality."

- 15 Because of its IR focus, emphasis will be placed on contributions to special editions of *International Studies Quarterly* (1989, 1990). The studies by George (1989), Der Derian (1990), George and Campbell (1990), Lapid (1989), Holsti (1989), Biersteker (1989), Ashley and Walker (1990) provide a good sense of priorities among postmodern critics of IR.
- 16 Alger seems to be confusing normative and scientific issues here. The validity of a theory has absolutely nothing to do with the research topic; it is a function of the logical consistency of a theory's assumptions and the theory's empirical strength. The relative importance of various research areas is a normative issue that lies outside the scientific enterprise.
- 17 For example, Winch (1972) and Cox (1986, 1992) apply many of these assertions to the "rationality" assumption, arguing that even rationality must be interpreted within a cultural context.
- 18 Indeed, a great deal of Soviet behaviour from 1917 onward revealed the same realist pattern as did the behaviour of non-Communist states. Brecher (1993, chaps. 2–5) found, for example, that, over the past half century, the behavioural patterns of the United States and the Soviet Union, particularly during international crises, were very similar, not just comparable.
- 19 According to Lynch (36–7), Soviet studies of international relations moved away "from rigid determinism of orthodox Soviet political economy (pre-1956) ... toward conceptions that accord a greater importance to such categories as politics, the state, and the very 'system' of international relations."
- 20 Most of what is acknowledged as "fact" is generally accepted and used freely, and any errors in generalization produced by cultural nuances are likely to be insignificant. As Nicholson (1992, 34) points out, "the degree to which facts depend on disputed conceptual frameworks for their meaning varies in such a way that there is a domain that is relatively unproblematic." For example, although scholars may have disagreements over how to operationalize a concept like "alliances" (whether on the basis of economic, political, or military agreements, a combination of all three, formal or informal arrangements, actual resolve to defend an ally, and so on), they can still evaluate a particular theory of alliance behaviour based on the logical consistency of the theory's central axioms, assumptions, and propositions. Once again, it is contingent upon postmodernists to go beyond arguments about the potential impact of these differences and demonstrate their actual effect.

- 21 Even “facts” and “empirical evidence” must be considered time-specific, in a certain sense. Nicholson’s (1992, 34) discussion of the telephone is a good illustration of the argument. The existence of the telephone depends on the acceptance of a particular conceptual framework, in this case one that is apparent to most people in the twentieth century but not to someone in the sixteenth century, who would have had “no sets of concepts to recognize it or relate to it ... no conceptual framework into which it fits ... Facts are only understood in some sort of context, and are not free-floating, simply waiting to be observed ... It is only when one gets to the point of being able to say that one sees the object ‘as a telephone’ that it makes sense and that a ‘fact’ such as ‘x percent of households in Britain have telephones’ becomes meaningful” (1992, 34). Of course, although similar conceptual frameworks would not exist, the empirical fact would still remain; the only difference would be that description of the object would not be as precise.
- 22 For example, it is conceivable that nuclear proliferation could reestablish a multipolar international system. The extensive theoretical work on conflict under multipolar conditions might become increasingly relevant once again.
- 23 Indeed, because change is inevitable, many scholars in the realist tradition continue to assume the existence of specified international conditions, even though, as the critics claim, they may only be approximations of current reality. The very process of theorizing requires nothing less.

Bibliography

- Achen, Christopher H., and Duncan Snidal. 1989. "Rational Deterrence Theory and Comparative Case Studies." *World Politics* 41: 143-69.
- Alger, Chadwick F. 1976. "The Researcher in the United Nations: Evolution of a Research Strategy." In *In Search of Global Patterns*, edited by J.N. Rosenau, 58-72. New York: Free Press.
- Allison, Graham T. 1971. *Essence of Decision: Explaining the Cuban Missile Crisis*. Boston: Little, Brown & Co.
- Art, Robert. 1985. "Bureaucratic Politics and American Foreign Policy: A Critique." In *International Politics: Anarchy, Force, Political Economy, and Decision Making* 2d ed., edited by Robert Art and Robert Jervis, 467-90. Boston: Little, Brown & Co.
- Ashley, R.K. 1986. "The Poverty of Neorealism." In *Neorealism and its Critics*, edited by Robert Keohane, 255-300. New York: Columbia University Press.
- Ashley, R.K., and R.B.J. Walker. 1990. "Speaking the Language of the Exile: Dissident Thought in International Studies." *International Studies Quarterly* 34: 259-68.
- Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.
- Axelrod, Robert, and Robert O. Keohane. 1985. "Achieving Cooperation under Anarchy." *World Politics* 38: 226-54.
- Azar, Edward E. 1986. *International Conflict: Theory and Practice*. Sussex, England: Wheatsheaf.
- Banks, Michael. 1990. "The International Relations Discipline: Asset or Liability for Conflict Resolution." In *Conflict: Readings in Management and Resolution*, edited by John Burton and Frank Dukes, 275-87. New York: St Martin's Press.

- Beer, Francis A. 1986. "Games and Metaphors." *Journal of Conflict Resolution* 30: 171-91.
- Behr, Roy L. 1981. "Nice Guys Finish Last - Sometimes." *Journal of Conflict Resolution* 25: 289-300.
- Berkowitz, B.D. 1985. "Proliferation, Deterrence, and the Likelihood of Nuclear War." *Journal of Conflict Resolution* 29: 112-36.
- Betts, Richard K. 1987. *Nuclear Blackmail and Nuclear Balance*. Washington, DC: The Brookings Institution.
- Biersteker, Thomas J. 1989. "Critical Reflections on Post-Positivism in International Relations." *International Studies Quarterly* 33: 263-7.
- Blechman, Barry M., Stephen S. Kaplan, et al. 1978. *Force without War*. Washington, DC: The Brookings Institution.
- Bobrow, D.B. 1982. "Uncoordinated Giants." In *Foreign Policy: U.S.A./USSR*, edited by C.W. Kegley and P. McGowen, 23-30. Beverly Hills, CA: Sage.
- Brams, Steven J. 1975. *Game Theory and Politics*. New York: Macmillan.
- 1985. *Superpower Games: Applying Game Theory to Superpower Conflict*. New Haven, CT: Yale University Press.
- Brams, Steven J., and D. Marc Kilgour. 1987a. "Threat Escalation and Crisis Stability: A Game-Theoretic Analysis." *American Political Science Review* 81: 833-50.
- 1987b. "Optimal Threats." *Operations Research* 35: 524-36.
- 1988. *Game Theory and National Security*. New York: Basil Blackwell.
- Brams, Steven J., and Donald Wittman. 1981. "Nonmyopic Equilibrium in 2x2 Games." *Conflict Management and Peace Science* 6: 39-62.
- Braybrooke, David, and Charles E. Lindblom. 1963. *A Strategy of Decision*. New York: Free Press.
- Brecher, Michael. 1973. *The Foreign Policy System of Israel*. New Haven, CT: Yale University Press.
- 1975. *Decisions in Israel's Foreign Policy*. New Haven, CT: Yale University Press.
- 1977. "Toward a Theory of International Crisis Behavior: A Preliminary Report." *International Studies Quarterly* 21: 39-74.
- 1993. *Crises in World Politics: Theory and Reality*. Oxford: Pergamon Press.
- Brecher, Michael, and Benjamin Geist. 1980. *Decisions in Crisis: Israel 1967 and 1973*. Berkeley, CA: University of California Press.
- Brecher, Michael, and Patrick James. 1986. *Crisis and Change in World Politics*. Boulder, CO: Westview Press.
- 1988. "Stability and Polarity: New Paths to Inquiry." *Journal of Peace Research* 25: 31-42.
- Brecher, Michael, Jonathan Wilkenfeld, et al. 1988. *Crises in the Twentieth Century, Volume I: Handbook of International Crises*. Oxford: Pergamon Press.
- 1989. *Crisis, Conflict, and Instability*. Oxford: Pergamon Press.
- 1991. "International Crises and Global Instability: The Myth of the 'Long Peace.'" In *The Long Postwar Peace*, edited by Charles Kegley, 85-104. New York: Harper Collins.

- Bremer, S.A. 1976. "Obstacles to the Accumulation of Knowledge." In *In Search of Global Patterns*, edited by J.N. Rosenau, 204-8. New York: Free Press.
- Bremer, S.A., J.D. Singer, and U. Luterbacher. 1973. "The Population Density and War Proneness of European Nations, 1816-1965." *Comparative Political Studies* 6: 329-48.
- British Broadcasting Corporation. 1975. *Summary of World Broadcasting, Part IV: The Middle East*: 5074/A/9, December 2.
- Brodie, Bernard. 1959a. *Strategy in the Missile Age*. Princeton, NJ: Princeton University Press.
- 1959b. "The Anatomy of Deterrence." *World Politics* 11: 173-91.
- 1973. *From Crossbow to H-Bomb*. Bloomington, IN: Indiana University Press.
- ed. 1946. *The Absolute Weapon*. New York: Harcourt Brace.
- Bueno de Mesquita, Bruce. 1980a. "An Expected Utility Theory of International Conflict." *American Political Science Review* 74: 917-31.
- 1980b. "Theories of International Conflict: An Analysis and an Appraisal." In *Handbook of Conflict: Theory and Research*, edited by T.R. Gurr, 367-98. New York: Macmillan.
- 1981. *The War Trap*. New Haven, CT: Yale University Press.
- 1985. "Toward a Scientific Understanding of International Conflict." *International Studies Quarterly* 29: 121-36.
- Bueno de Mesquita, Bruce, and David Lalman. 1986. "Reason and War." *American Political Science Review* 80: 1113-29.
- 1992. *War and Reason*. New Haven, CT: Yale University Press.
- Bueno de Mesquita, Bruce, and William Riker. 1982. "Assessing the Merits of Nuclear Proliferation." *Journal of Conflict Resolution* 26: 283-306.
- Bull, Hedley. 1966. "International Theory: The Case for the Classical Approach." *World Politics* 18: 361-77.
- Butterworth, R.L. 1976. *Managing Interstate Conflict, 1945-1974: Data with Synopsis*. Pittsburgh, PA: University of Pittsburgh Press.
- Choucri, N., and C. North. 1972. "Alternative Dynamics of International Conflicts: Population, Resources, Technology, and Some Implications for Policy." *World Politics* 24: 80-122.
- Cimbala, Stephen J. 1988. *Nuclear Strategizing: Deterrence and Reality*. New York: Praeger.
- Gioffi-Revilla, Claudio. 1983. "A Probability Model of Credibility." *Journal of Conflict Resolution* 27: 73-108.
- Clancy, Tom, and Russell Seitz. 1992. "Five Minutes Past Midnight: And Welcome to the Age of Proliferation." *The National Interest* 26 (winter): 3-12.
- Clark, Mary E. 1990. "What is Science For? Reintroducing Philosophy into the Political Science Classroom." In *Conflict: Readings in Management and Resolution*, edited by John Burton and Frank Dukes, 51-70. New York: St Martin's Press.
- Cohen, Raymond. 1979. *Threat Perception in International Crises*. Madison, WI: University of Wisconsin Press.

- Corson, W.H. II. 1970. "Conflict and Cooperation in East-West Crises: Dynamics of Crisis Interaction." Paper presented at the annual meeting of the International Studies Association, Washington, DC, 13 April.
- Cox, Robert. 1986. "Social Forces, States and World Order: Beyond International Relations Theory." In *Neorealism and its Critics*, edited by Robert Keohane, 204-54. New York: Columbia University Press.
- 1992. "Towards a Post-Hegemonic Conceptualization of World Order: Reflections on the Relevancy of Ibn Khaladin." In *Governance without Government: Order and Change in World Politics*, edited by J.N. Rosenau and Ernst-Otto Czempiel, 132-59. Cambridge: Cambridge University Press.
- Craig, Gordon A., and Alexander George. 1990. *Force and Statecraft: Diplomatic Problems of Our Time*. New York: Oxford University Press.
- Crawley, James B. 1966. *Japan's Quest for Autonomy*. Princeton: Princeton University Press.
- Dacey, Raymond, and Norman Pendegrift. 1988. "The Optimality of Tit-For-Tat." *International Interaction* 15: 45-64.
- Dawisha, Adeed. 1978. "Syrian Intervention in Lebanon, 1975-1976." *Jerusalem Journal of International Relations* 3: 245-63.
- Dawisha, Karen. 1984. *The Kremlin and the Prague Spring: Decisions in Crisis*. Berkeley, CA: University of California Press.
- Deeb, Marius. 1980. *The Lebanese Civil War*. New York: Praeger.
- Der Derian, James. 1990. "The (S)pace of International Relations: Simulation, Surveillance, and Speed." *International Studies Quarterly* 34: 295-310.
- Deutsch, Karl. 1968. "A Critique of Deterrence Theory." In *International Politics: Anarchy, Force, Imperialism*, edited by Robert J. Art and Robert Jervis, 240-4. Boston: Little, Brown & Co.
- 1988. *The Analysis of International Relations*. 3d ed. Englewood Cliffs, NJ: Prentice-Hall.
- Deutsch, Karl, and David Singer. 1964. "Multipolar Power Systems and International Stability." *World Politics* 21: 390-406.
- Dingman, Roger. 1988. "Atomic Diplomacy during the Korean War." *International Security* 13: 50-91.
- Doran, Charles F. 1973. "The Theory of Bounded Deterrence." *Journal of Conflict Resolution* 17: 243-69.
- 1989. "Power Cycle Theory of System Structure and Stability." In *Handbook of War Studies*, edited by Manus Midlarsky, 83-110. Boston: Unwin Hyman.
- Dougherty, James E., and Robert L. Pfaltzgraff Jr. 1990. *Contending Theories of International Relations: A Comprehensive Survey*. New York: Harper and Row.
- Dowty, Alan. 1984. *Middle East Crisis: U.S. Decision Making in 1958, 1970, and 1973*. Berkeley, CA: University of California Press.
- Dukes, Frank. 1990. "Action Research." In *Conflict: Readings in Management and Resolution*, edited by John Burton and Frank Dukes, 288-98. New York: St Martin's Press.

- Dunn, David J. 1981. "The Emergence of Change as a Theoretical Concept in International Relations." In *Change and the Study of International Relations: The Evaded Dimension*, edited by Barry Buzan and R.J.B. Jones, 71–84. London: Francis Pinter.
- Ellsberg, Daniel. 1961. "The Crude Analysis of Strategic Choice." *American Economic Review Papers* 51: 472–8.
- 1992. "Nuclear Security and the Soviet Collapse." Interview by Jerry Sanders and Richard Caplan. *World Policy Journal* 28: 135–56.
- Evron, Yair. 1987. *War and Intervention in Lebanon: The Israeli-Syrian Deterrence Dialogue*. London: Croom Helm.
- Farnham, B. 1992. "Roosevelt and the Munich Crisis: Insights from Prospect Theory." *Political Psychology* 13: 205–35.
- Fearon, James. 1994. "Signaling versus the Balance of Power and Interests: An Empirical Test of a Crisis Bargaining Model." *Journal of Conflict Resolution* 38: 236–69.
- Feaver, Peter. 1993. "Command and Control in Emerging Nuclear Nations." *International Security* 17(3): 160–87.
- 1994. "Proliferation Optimism and Theories of Nuclear Operations." In *The Proliferation Puzzle: Why Nuclear Weapons Spread and What Results*, edited by Zachary Davis and Benjamin Frankel, 159–91. London: Frank Cass and Company.
- Feldman, Shai. 1981. "A Nuclear Middle East." *Survival* 23 (May-June): 107–66.
- Festinger, Leon. 1957. *A Theory of Cognitive Dissonance*. Stanford, CA: Stanford University Press.
- 1964. *Conflict, Decision and Dissonance*. Stanford, CA: Stanford University Press.
- Feyerabend, Paul. 1975. *Against Method*. London: New Left Books.
- Fink, Clinton F. 1965. "More Calculations about Deterrence." *Journal of Conflict Resolution* 9: 54–65.
- Fischhoff, Baruch. 1987. "Do We Want a Better Theory of Deterrence?" *Journal of Social Issues* 43: 73–7.
- Fisher, Roger. 1964. "Fractionating Conflict." In *International Conflict and Behavioral Science: The Craienville Papers*, edited by R. Fisher, 91–109. New York: Basic Books.
- Fleming, Paula. 1991. "Unstable Ground: The Rationality Critique of Deterrence Theory." Paper presented at the annual meeting of the American Political Science Association, Washington, DC.
- Foot, Rosemary J. 1989. "Nuclear Coercion and the Ending of the Korean Conflict." *International Security* 13: 92–112.
- Freedman, Lawrence. 1986. "Nuclear Weapons and Strategy." In *Studies of War and Peace*, edited by Oyvind Osterud, 71–89. Oslo: Norwegian University Press.

- Freedman, Robert O. 1975. *Soviet Policy toward the Middle East since 1970*. New York: Praeger.
- 1987. “The Soviet Union and the Civil War in Lebanon, 1975–1976.” *The Jerusalem Journal of International Relations* 3: 60–93.
- Freud, Sigmund. 1938. *The Basic Writings of Sigmund Freud*. New York: Modern Library.
- 1950. “Why War?” In *Readings in World Politics*, edited by Robert A. Goldwin, Ralph Lerner, et al., 19–32. New York: Oxford University Press.
- Friedman, James W. 1986. *Game Theory with Applications to Economics*. Oxford: Oxford University Press.
- Fukuyama, Francis. 1989. “The End of History?” *The National Interest* 16: 3–16.
- Gaddis, John Lewis. 1986. “The Long Peace: Elements of Stability in the Postwar International System.” *International Security* 10: 92–142.
- 1987. *The Long Peace: Inquiries into the History of the Cold War*. New York: Oxford University Press.
- 1991. “Great Illusions, the Long Peace, and the Future of the International System.” In *The Long Postwar Peace: Contending Explanations and Predictions*, edited by Charles Kegley, 25–51. New York: Harper Collins.
- Garthoff, R.L. 1962. *Soviet Strategy in the Missile Age*. New York: Praeger.
- Gauthier, D. 1984. “Deterrence, Maximization, and Rationality.” In *The Security Gamble*, edited by D. Maclean, 474–95. New Jersey: Rowman & Allanheld.
- Geller, Daniel S. 1990. “Nuclear Weapons, Deterrence, and Crisis Escalation.” *Journal of Conflict Resolution* 34: 291–310.
- George, Alexander L. 1985. “Adapting to Constraints on Rational Decision Making.” In *International Politics: Anarchy, Force, Political Economy, and Decision Making*, edited by Robert Art and Robert Jervis, 491–509. Boston: Little, Brown & Co.
- 1986. “Problems of Crisis Management and Crisis Avoidance in U.S.-Soviet Relations.” In *Studies of War and Peace*, edited by Oyvind Osterud, 202–26. Oslo: Norwegian University Press.
- ed. 1991. *Avoiding War: Problems of Crisis Management*. Boulder, CO: Westview Press.
- George, Alexander L., and Richard Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.
- 1989. “Deterrence and Foreign Policy.” *World Politics* 41: 170–82.
- George, Jim. 1989. “International Relations and the Search for Thinking Space: Another View of the Third Debate.” *International Studies Quarterly* 33: 269–79.
- George, Jim, and David Campbell. 1990. “Patterns of Dissent and the Celebration of Differences: Critical Social Theory and International Relations.” *International Studies Quarterly* 34: 269–93.
- Geortz, Gary. 1995. “(Enduring) Rivalries and the Study of Deterrence.” Paper presented at conference, “Deterrence in Enduring Rivalries,” Georgetown University, Washington, DC, 8–9 March, 1996.

- Geertz, Gary, and Bear F. Braumoelle. 1996. "The Necessary and Sufficient Condition Hypothesis." Paper presented at the annual meeting of the International Studies Association, San Diego, California, 16–20 April.
- Giddens, A. 1979. *Central Problems in Social Theory*. London: Macmillan.
- Gillespie, John V., and Dina Zinnes. 1975. "Progression in Mathematical Models of International Relations." *Synthese* 31: 298–321.
- Gilpin, R.G. 1986. "The Richness of the Tradition of Political Realism." In *Neorealism and Its Critics*, edited by Robert Keohane, 301–21. New York: Columbia University Press.
- Glaser, Charles L. 1990. *Analysing Strategic Nuclear Policy*. Princeton, NJ: Princeton University Press.
- Gochman, Charles S. 1975. "Status, Conflict, and War: The Major Powers, 1820–1970." PhD diss., University of Michigan.
- Golan, Galia. 1977. *Yom Kippur and After*. Cambridge: Cambridge University Press.
- Goldstein, S.J. 1991. "Reciprocity in Superpower Relations: An Empirical Analysis." *International Studies Quarterly* 35: 195–209.
- Goldstein, S.J., and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.
- Haas, Michael. 1965. "Societal Approaches to the Study of War." *Journal of Peace Research* 2: 307–23.
- 1986. "Research on International Crises." *International Interactions* 13: 23–58.
- Hackett, Robert. 1993. "Structuring the Hierarchy of Dissent." *Peace Review* 5(1): 19–26.
- Haley, Edward, and Lewis Snider. eds. 1979. *Lebanon in Crisis: Participants and Issues*. Syracuse, NY: Syracuse University Press.
- Halperin, Morton H., and Arnold Kaner. 1985. "The Bureaucratic Perspective." In *International Politics: Anarchy, Force, Political Economy, and Decision Making*. 2d ed., edited by Robert Art and Robert Jervis, 439–68. Boston: Little, Brown & Co.
- Harvey, Frank. 1995. "Rational Deterrence Theory Revisited: A Progress Report." *Canadian Journal of Political Science* 28: 403–36.
- 1996a. "Rigor Mortis or Rigor, More Tests: Deterrence and Compellence in Protracted Crises." Paper presented at the annual meeting of the International Studies Association. San Diego, California, 16–20 April.
- 1996b. "Deterrence and Compellence in Protracted Crises: The Case of Bosnia-Herzegovina, 1993–1995." *Security Studies*. Forthcoming.
- 1997. "Deterrence and Ethnic Conflict: The Case of Bosnia-Herzegovina, 1993–94." *Security Studies*. Forthcoming, winter 1996–97.
- Harvey, Frank, and Patrick James. 1992. "Nuclear Deterrence Theory: The Record of Aggregate Testing and an Alternative Research Agenda." *Conflict Management and Peace Science* 12: 17–45.
- Herek, G.M., I.L. Janis, and Paul Huth. 1987. "Decision Making during International Crises. Is Quality of Process Related to Outcome?" *Journal of Conflict Resolution* 30: 497–531.

- 1989. “Quality of U.S. Decision Making during the Cuban Missile Crisis. Major Errors in Welch’s Reassessment.” *Journal of Conflict Resolution* 33: 446–59.
- Hermann, C.F. 1972. “Policy Classification: A Key to the Comparative Study of Foreign Policy.” In *The Analysis of International Politics*, edited by J.N. Rosenau, 58–79. New York: Free Press.
- Herz, John H. 1951. *Political Realism and Political Idealism*. Chicago: University of Chicago Press.
- Hoffmann, Stanley. 1990. “Correspondence.” *International Security* 15: 191–9.
- Holsti, Kal. 1985. *The Dividing Discipline: Hegemony and Diversity in International Theory*. Boston: Allen & Unwin.
- 1989. “Mirror, Mirror on the Wall, Which Are the Fairest Theories of All?” *International Studies Quarterly* 33: 255–61.
- Holsti, O.R. 1965. “The 1914 Case.” *American Political Science Review* 59: 365–78.
- 1971. “Crisis Stress and Decision Making.” *International Social Science Journal* 22: 55–67.
- 1972. *Crisis, Escalation, War*. Montreal: McGill-Queen’s University Press.
- 1976. “Foreign Policy Decision Makers Viewed Psychologically: ‘Cognitive Process’ Approaches.” In *In Search of Global Patterns*, edited by J.N. Rosenau, 120–44. New York: Free Press.
- Holsti, O.R., and Alexander George. 1975. “The Effects of Stress on the Performance of Foreign Policy Makers.” In *Political Science Annual: An International Review*, edited by C.P. Cotter, 255–319. Indianapolis: Bobbs-Merrill.
- Holsti, O.R., R.C. North, and R. Brody. 1968. “Perceptions and Actions in the 1914 Crisis.” In *Quantitative International Politics: Insights and Evidence*, edited by J.D. Singer, 123–58. New York: Free Press.
- Hopple, G.W., and P.J. Rossa. 1981. “International Crisis Analysis: Recent Developments and Future Directions.” In *Cumulation in International Relations Research*, edited by T. Hopmann, D. Zinnes, and J.D. Singer, 65–97. Denver, CO: University of Denver.
- Huth, Paul K. 1988a. *Extended Deterrence and the Prevention of War*. New Haven, CT: Yale University Press.
- 1988b. “Extended Deterrence and the Outbreak of War.” *American Political Science Review* 82: 423–43.
- 1990. “The Extended Deterrent Value of Nuclear Weapons.” *Journal of Conflict Resolution* 34: 270–90.
- Huth, Paul K., and Bruce M. Russett. 1984. “What Makes Deterrence Work? Cases from 1900–1980.” *World Politics* 36: 496–526.
- 1988. “Deterrence Failure and Crisis Escalation.” *International Studies Quarterly* 32: 29–45.
- 1990. “Testing Deterrence Theory: Rigor Makes a Difference.” *World Politics* 42: 466–501.

- 1993. "General Deterrence between Enduring Rivalries." *American Political Science Review* 87: 61-73.
- Huth, Paul K., Christopher Gelpi, and Scott Bennett. 1993. "The Escalation of Great Power Militarized Disputes: Testing Rational Deterrence Theory and Structural Realism." *American Political Science Review* 87: 609-38.
- Jabber, F. 1971. *Israel and Nuclear Weapons*. London: International Institute for Strategic Studies.
- James, Patrick. 1988. *Crisis and War*. Montreal: McGill-Queen's University Press.
- 1991. "Rational Retaliation: Superpower Crises, 1948-1979." *Public Choice* 68: 117-35.
- 1993. "Structural Realism as a Research Enterprise: Toward Elaborated Structural Realism." *International Political Science Review* 14: 123-48.
- James, Patrick, and Frank Harvey. 1989. "Threat Escalation and Crisis Stability: Superpower Cases, 1948-1979." *Canadian Journal of Political Science* 22: 523-45.
- 1992. "The Most Dangerous Game: Superpower Rivalry in International Crises, 1948-1985." *Journal of Politics* 54: 25-51.
- Janis, Irving. 1982. *Groupthink*. 2d ed. Boston: Houghton Mifflin.
- 1989. *Crucial Decisions: Leadership in Policymaking and Crisis Management*. New York: Free Press.
- Janis, Irving, and Leon Mann. 1977. *Decision Making*. New York: Free Press.
- Jervis, Robert. 1976a. *Perception and Misperception in International Politics*. Princeton, NJ: Princeton University Press.
- 1976b. "Cumulation, Correlations and Woozles." In *In Search of Global Patterns*, edited by J.N. Rosenau, 181-5. New York: Free Press.
- 1978. "Cooperation under the Security Dilemma." *World Politics* 30: 167-214.
- 1979. "Deterrence Theory Revisited." *World Politics* 31: 289-324.
- 1980. "Why Nuclear Superiority Doesn't Matter." *Political Science Quarterly* 94: 617-33.
- 1985. "Symposium on Methodological Foundations of the Study of International Conflict." *International Studies Quarterly* 29: 121-54.
- 1989. "Rational Deterrence: Theory and Evidence." *World Politics* 41: 183-207.
- 1992. "Political Implications of Loss Aversion." *Political Psychology* 13(2): 187-204.
- Jervis, Robert, Richard N. Lebow, and Janice G. Stein. 1985. *Psychology and Deterrence*. Baltimore, MD: Johns Hopkins University Press.
- Jukes, Geoffrey. 1985. *Hitler's Stalingrad Decisions*. Berkeley, CA: University of California Press.
- Kahana, Ephraim. 1994. "Towards a Nuclear Middle East - The Israeli Dilemma: Proliferation, Strategy and Arms Control." Paper presented at the annual meeting of the International Studies Association, Washington, DC.

- Kahn, Herman. 1961. *On Thermonuclear War*. Princeton, NJ: Princeton University Press.
- 1962. *Thinking about the Unthinkable*. New York: Horizon.
- Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2): 263–92.
- Kaplan, Morton A. 1958. "The Calculus of Nuclear Deterrence." *World Politics* 11: 20–43.
- 1966. "The New Great Debate: Traditionalism vs. Science in International Relations." *World Politics* 19: 1–20.
- Kegley, Charles W., ed. 1991. *The Long Postwar Peace: Contending Explanations and Predictions*. New York: Harper Collins.
- Khalidi, Walid. 1979. *Conflict and Violence in Lebanon: Confrontation in the Middle East*. Harvard Studies in International Affairs, no. 38. Cambridge: Harvard University.
- Kennedy, Paul. 1988. *The Rise and Fall of Great Powers: Economic Change and Military Conflict 1500–2000*. New York: Random House.
- Kenny, Anthony. 1985. *The Logic of Deterrence*. Chicago: University of Chicago Press.
- Keohane, Robert. 1983. "Theory of World Politics: Structural Realism and Beyond." In *Political Science: The State of the Discipline*, edited by A. Finifter. Washington, DC: The American Political Science Association.
- 1984. *After Hegemony: Cooperation and Discord in World Political Economy*. Princeton, NJ: Princeton University Press.
- 1986a. "Theory of World Politics." In *Neorealism and its Critics*, edited by Robert Keohane, 158–203. New York: Columbia University Press.
- ed. 1986b. *Neorealism and Its Critics*. New York: Columbia University Press.
- 1990. "Correspondence." *International Security* 15: 191–9.
- Keohane, Robert, and Joseph Nye. 1977. *Power and Interdependence: World Politics in Transition*. Boston: Little, Brown & Co.
- Kissinger, Henry. 1982. *Years of Upheaval*. Boston: Little, Brown & Co.
- Kolodziej, Edward. 1987. "The Limits of Deterrence Theory." *Journal of Social Issues* 43: 123–33.
- Krasner, Stephen D. 1985. "Symposium on Methodological Foundations of the Study of International Conflict." *International Studies Quarterly* 29: 121–54.
- Krell, Gert. 1987. "Groping Towards Synthesis? A Fresh Approach to Deterrence Theory." *Journal of Social Issues* 43: 111–14.
- Kugler, Jacek. 1984. "Terror Without Deterrence." *Journal of Conflict Resolution* 28: 470–506.
- Kupperman, R.H., and R.A. Smith. 1971. "Some Computational Techniques and Theorems Concerning Mutual Deterrence Relationships." *Journal of Conflict Resolution* 1: 4–31.
- Kuhn, T.S. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

- Lapid, Yosef. 1989. "The Third Debate: On the Prospects of International Theory in a Post-Positivist Era." *International Studies Quarterly* 33: 235-54.
- Larus, Joel. 1967. *Nuclear Weapons Safety and the Common Defence*. Columbus, OH: Ohio State University Press.
- Lawson, Fred. 1984. "Syria's Intervention in the Lebanese Civil War." *International Organization* 38: 451-77.
- Lebow, Richard N. 1981. *Between Peace and War: The Nature of International Crises*. Baltimore, MD: The Johns Hopkins University Press.
- Lebow, Richard N., and Janice G. Stein. 1987. "Beyond Deterrence." *Journal of Social Issues* 43: 5-71.
- 1989a. "Rational Deterrence Theory: I Think, Therefore I Deter." *World Politics* 41: 208-24.
 - 1989b. "When Does Deterrence Succeed and How Do We Know?" Paper presented at the annual meeting of the International Studies Association, Washington, DC.
 - 1990. "Deterrence: The Elusive Dependent Variable." *World Politics* 42: 336-69.
- Leng, Russell J. 1988. "Crisis Learning Games." *American Political Science Review* 82: 179-94.
- 1990. "Influence Techniques in Interstate Crises: Realism vs. Reciprocity." Paper presented at the annual meeting of the International Studies Association, Washington, DC.
- Leng, Russell J., and Hugh G. Wheeler. 1979. "Influence Strategies, Success and War." *Journal of Conflict Resolution* 23: 655-84.
- Levy, Jack. 1985. "The Polarity of the System and International Stability: An Empirical Analysis." In *Polarity and War: The Changing Structure of International Conflict*, edited by Alan Ned Sabrosky, 41-66. Boulder, CO: Westview Press.
- 1988. "When Do Deterrent Threats Work?" *British Journal of Political Science* 18: 485-512.
 - 1989a. "The Causes of War: A Review of Theories and Evidence." In *Behavior, Society, and Nuclear War*, edited by Philip E. Tetlock et al., 209-333. London: Oxford University Press.
 - 1989b. "The Diversionary Theory of War: A Critique." In *Handbook of War Studies*, edited by Manus Midlarsky, 259-88. Boston: Unwin Hyman.
 - 1992a. "An Introduction to Prospect Theory." *Political Psychology* 13(2): 171-86.
 - 1992b. "Prospect Theory and International Relations: Theoretical Applications and Analytical Problems." *Political Psychology* 13(2): 283-310.
- Lieberman, Elli. 1995. "What Makes Deterrence Work?" Lessons from the Egyptian-Israeli Enduring Rivalry." *Security Studies* 4: 851-910.
- Lindskold, S., Brian Betz, and Pamela Walters. 1986. "Transforming Competitive or Cooperative Climates." *Journal of Conflict Resolution* 30: 99-114.

- Lynch, Allen. 1989. *The Soviet Study of International Relations*. Cambridge: Cambridge University Press.
- MacInnis, Major-General J.A. 1995. Letter to the author, 24 November 1994.
- Ma'oz, Moshe. 1988. *Asad: The Sphinx of Damascus*. London: Weidenfeld and Nicolson.
- Ma'oz, Zeev. 1988. "Conflict Datasets: Definitions and Measurements." *International Interactions* 14: 165-71.
- Martin, Richard. 1976. In *In Search of Global Patterns*, edited by James Rosenau, 212-15. New York: Free Press.
- Maxwell, S. 1968. *Rationality in Deterrence*. Adelphi Papers no. 50. London: International Institute for Strategic Studies.
- McClelland, Charles. 1961. "The Acute International Crisis." *World Politics* 14: 182-204.
- McGinnis, Michael. 1988. "Domestic Political Competition and the Unitary Rational Actor Assumption: Models of Security Rivalries." Paper presented at the annual meeting of the International Studies Association, St Louis, Missouri.
- Mearsheimer, John J. 1990a. "Back to the Future: Instability in Europe after the Cold War." *International Security* 15: 5-56.
- 1990b. "Correspondence." *International Security* 15: 191-9.
- 1990c. "Correspondence." *International Security* 15: 216-22.
- 1993. "The Case for a Ukrainian Nuclear Deterrent." *Foreign Affairs* 72(3): 50-66.
- Meo, Leilla. 1977. "Syria and the Conflict in Lebanon." In *Lebanon: Crisis and Challenge in the Arab World*, edited by Fouad Mougharabi and Naseer Aruri, 26-31. Special report no. 1. Detroit, MI: Association of Arab-American University Graduates.
- Michener, Andrew H., K. Yuen, and S. Geisheker. 1980. "Nonsymmetry and Core Size in n -Person Sidepayment Games." *Journal of Conflict Resolution* 24: 495-523.
- Michener, Andrew H., and Kathryn Potter. 1981. "Generalizability of Tests in n -Person Sidepayment Games." *Journal of Conflict Resolution* 25: 733-49.
- Milburn, T.W. 1959. "What Constitutes Effective Deterrence?" *Journal of Conflict Resolution* 3: 138-44.
- Modelski, George, and William Thompson. 1989. "Long cycles and Global War." In *Handbook of War Studies*, edited by Manus Midlarsky, 23-54. Boston: Unwin Hyman.
- Molander, Roger C., and Peter A. Wilson. 1994. "On Dealing with the Prospect of Nuclear Chaos." *The Washington Quarterly* 17(3): 19-39.
- Morgan, Patrick M. 1977. *Deterrence: A Conceptual Analysis*. Beverly Hills, CA: Sage.
- Morgan, T. Clifton. 1984. "A Spatial Model of Crisis Bargaining." *International Studies Quarterly* 28: 407-26.

- Morgenthau, Hans. 1960. *Politics among Nations*. 3d ed. New York: Alfred A. Knopf.
- Mowlana, Hamid. 1993. "Triumph of the Image." *Peace Review* 5(1): 27-34.
- Mueller, John. 1988. "The Essential Irrelevance of Nuclear Weapons: Stability in the Post War World." *International Security* 13: 55-79.
- Mueller, Ulrich. 1987. "Optimal Retaliation for Optimal Cooperation." *Journal of Conflict Resolution* 31: 692-724.
- Naroll, R. 1969. "Deterrence in History." In *Theory and Research on the Causes of War*, edited by D.G. Pruitt and R.C. Snyder, 150-64. Englewood Cliffs, NJ: Prentice Hall.
- Naroll, R., V.L. Bullough, and F. Narrol. 1974. *Military Deterrence in History*. Albany, NY: State University of New York Press.
- Nicholson, Michael. 1992. *Rationality and the Analysis of International Conflict*. Cambridge: Cambridge University Press.
- Nixon, Richard. 1978. *The Memoirs of Richard Nixon*. New York: Grosset & Dunlap.
- Noble, Paul. 1983. "The Arab System: Opportunities, Constraints and Pressures." In *The Foreign Policies of Arab States*, edited by B. Korany and A. Dessouki, 50-60. Boulder, CO: Westview Press.
- Oneal, John R. 1988. "The Rationality of Decision Making During International Crises." *Polity* 20: 598-622.
- Organski, A.F.K., and Jacek Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.
- Orme, John. 1987. "Deterrence Failures: A Second Look." *International Security* 11: 96-124.
- Osgood, Charles E. 1962. *An Alternative to War or Surrender*. Urbana, ILL: University of Illinois Press.
- 1979. "GRIT for MFBR: A Proposal for Unfreezing Force-Level Postures in Europe." *Peace Research Review* 8: 77-92.
- Oye, Kenneth A. 1985. "Explaining Cooperation under Anarchy: Hypotheses and Strategies." *World Politics* 38: 1-24.
- Pages, Erik R. 1991. "The Evolution of Deterrence Theory: A Review of the Literature." *International Studies* 16: 60-5.
- Paige, Glenn. 1968. *The Korean Decision: June 24-30*. New York: Free Press of Glencoe.
- Patchen, Martin. 1987. "Strategies for Eliciting Cooperation from an Adversary: Laboratory and International Findings." *Journal of Conflict Resolution* 31: 164-86.
- 1988. *Resolving Disputes between Nations: Coercion or Conciliation?* Durham, NC: Duke University Press.
- 1990. "Conflict and Cooperation in American-Soviet Relations: What Have We Learned from Quantitative Research?" Paper presented at the annual meeting of the International Studies Association, Washington, DC.

- Peterson, W.J. 1986. "Deterrence and Compellence: A Critical Assessment of Conventional Wisdom." *International Studies Quarterly* 30: 269-94.
- Phillips, D.L. 1977. *Wittgenstein and Scientific Knowledge: A Sociological Perspective*. London: Macmillan.
- Platig, Raymond. 1966. *International Relations Research: Problems of Evaluation and Advancement*. Carnegie Endowment for International Peace. Santa Barbara, CA: Clio Press.
- Post, Jerrold M. 1991. "The Impact of Crisis-Induced Stress on Policy Makers." *Avoiding War: Problems of Crisis Management*, edited by Alexander George, 471-96. San Francisco: Westview Press.
- Powell, Robert. 1990. *Nuclear Deterrence Theory: The Search for Credibility*. New York: Cambridge University Press.
- Quandt, William B. 1977. *Decade of Decisions*. Berkeley, CA: University of California Press.
- Rabinovich, Itamar. 1984. *The War for Lebanon: 1970-1983*. London: Cornell University Press.
- Rapoport, Anatol. 1960. *Fights, Games and Debates*. Ann Arbor, MI: University of Michigan Press.
- Rapoport, Anatol, and Albert M. Chammah. 1965. *Prisoner's Dilemma*. Ann Arbor, MI: University of Michigan Press.
- Raser, J.R. 1966. "Theories of Deterrence: Past Progress and Future Needs." *Peace Research Reviews* 3: 297-327.
- Rasler, Karen. 1984. "A Comparative Analysis of Monadic and Dyadic Perspectives on the 1976 Lebanese Crisis." *International Studies Quarterly* 28: 427-46.
- Rejai, Mostafa, and Cynthia Enloe. 1969. "Nation-States and State-Nations." *International Studies Quarterly* 13: 140-58.
- Richardson, J.L. 1988. "New Insights on International Crises." *Review of International Studies* 14: 309-16.
- Risse-Kappen, Thomas. 1990. "Correspondence." *International Security* 15: 216-22.
- Rosecrance, R. 1976. "The Failures of Quantitative Analysis: Possible Causes and Cures." In *In Search of Global Patterns*, edited by J.N. Rosenau, 174-9. New York: Free Press.
- Rosen, Stephen. 1977. "A Stable System of Mutual Nuclear Deterrence in the Arab-Israeli Conflict." *American Political Science Review* 71: 1367-83.
- Rosenau, James N. 1966. "Pre-Theories and Theories of Foreign Policy." In *Approaches to Comparative and International Politics*, edited by R.B. Farell, 27-92. Evanston, IL: Northwestern University.
- 1971. *The Scientific Study of Foreign Policy*. New York: Free Press.
 - 1990. *Turbulence in World Politics: A Theory of Change and Continuity*. Princeton, NJ: Princeton University Press.
 - 1992. "Governance, Order, and Change in World Politics." In *Governance without Government: Order and Change in World Politics*, edited by James

- Rosenau and Ernst-Otto Czempiel, 1–29. Cambridge: Cambridge University Press.
- ed. 1976. *In Search of Global Patterns*. New York: Free Press.
- Rosenau, Pauline. 1992. *Post-Modernism and the Social Sciences: Insights, Inroads, and Intrusions*. Princeton, NJ: Princeton University Press.
- Ruggie, John Gerard. 1983. “Continuity and Transformation in World Politics: Toward a Neorealist Synthesis.” *World Politics* 35: 261–85.
- Rummel, R.J. 1965. “A Social Field Theory of Foreign Conflict.” *Peace Research Society Papers* 4: 131–50.
- 1968. “National Attributes and Foreign Conflict Behavior.” In *Quantitative International Politics*, edited by J.D. Singer, 187–214. New York: Macmillan.
- Russett, Bruce M. 1963. “The Calculus of Deterrence.” *Journal of Conflict Resolution* 7: 97–109.
- 1967. “Pearl Harbour: Deterrence Theory and Decision Theory.” *Journal of Peace Research* 2: 89–106.
- 1987. “Further beyond Deterrence.” *Journal of Social Issues* 43: 99–104.
- 1990. “Correspondence.” *International Security* 15: 216–22.
- Sabrosky, Alan Ned. 1985. *Polarity and War: The Changing Structure of International Conflict*. Boulder, CO: Westview Press.
- Safran, Nadav. 1981. *Israel, the Embattled Ally*. Cambridge: Harvard University Press.
- Sagan, Scott. 1985. “Nuclear Alerts and Crisis Management.” *International Security* 9(4), 99–139.
- 1994. “The Perils of Proliferation.” *International Security* 18: 66–107.
- Schelling, Thomas C. 1960. *The Strategy of Conflict*. Oxford: Oxford University Press.
- 1966. *Arms and Influence*. New Haven, CT: Yale University Press.
- Shafir, Eldar. 1992. “Prospect Theory and Political Analysis: A Psychological Perspective.” *Political Psychology* 13(2): 311–22.
- Shlaim, Avi. 1983. *The United States and the Berlin Blockade, 1948–1949*. Berkeley, CA: University of California Press.
- Shubik, Martin. 1986. “Cooperative Game Solutions.” *Journal of Conflict Resolution* 30: 63–76.
- Singer, David. 1969. “The Level of Analysis Problem.” *World Politics* 14: 77–92.
- 1976. “Tribal Sins on the QIP Reservation.” In *In Search of Global Patterns*, edited by J.N. Rosenau, 167–73. New York: Free Press.
- 1981. “Accounting for International War: The State of the Discipline.” *Journal of Peace Research* 18: 1–15.
- Singer, David, and M. Small. 1972. *The Wages of War, 1816–1965: A Statistical Handbook*. New York: John Wiley.
- Siverson, R.M. 1976. “Some Suggestions for Improving Cumulation.” In *In Search of Global Patterns*, edited by J.N. Rosenau, 198–203. New York: Free Press.

- Small, Melvin, and J. David Singer. 1982. *Resort to Arms: International and Civil Wars, 1816-1980*. Beverly Hills, CA: Sage.
- Smith, T.C. 1981. *Trojan Peace: Some Deterrence Propositions Tested*. Monograph Series in World Affairs. Denver, CO: University of Denver.
- Smoke, Richard. 1976. "Theory for and about Policy." In *In Search of Global Patterns*, edited by J.N. Rosenau, 185-91. New York: Free Press.
- Snidal, Duncan. 1985. "The Game Theory of International Politics." *World Politics* 38: 25-57.
- 1991. "Relative Gains and the Pattern of International Cooperation." *American Political Science Review* 85: 701-26.
- Snyder, G.H. 1960. "Deterrence and Power." *Journal of Conflict Resolution* 4: 163-78.
- 1961. *Deterrence and Defence*. Princeton, NJ: Princeton University Press.
- 1971. "Prisoner's Dilemma and Chicken Models in International Politics." *International Studies Quarterly* 15: 66-103.
- 1972. "Crisis Bargaining." In *International Crises: Insights from Behavioral Research*, edited by Charles F. Hermann, 419-45. New York: Free Press.
- Snyder, G.H., and Paul Diesing. 1977. *Conflict among Nations: Bargaining, Decision Making, and System Structure in International Crises*. Princeton, NJ: Princeton University Press.
- Snyder, G.H., H.W. Bruck, and Burton Sapin. 1962. "Decision Making as an Approach to the Study of International Politics." In *Foreign Policy Decision Making*, edited by G.H. Snyder, H.W. Bruck, and Burton Sapin. New York: Free Press.
- Sokolovsky, V.D. 1975. *Soviet Military Strategy*. Edited with analysis and commentary by Harriet Fast Scott. London: Macdonald and James.
- Stein, Janice Gross. 1985. "Calculation, Miscalculation and Conventional Deterrence I: The View from Cairo." In *Psychology and Deterrence*, edited by Robert Jervis, Richard Lebow, and Janice Stein, 60-88. Baltimore, MD: Johns Hopkins University Press.
- Stein, Janice Gross, and Louis Pauly, eds. 1993. *Choosing to Cooperate: How States Avoid Losses*. Baltimore, MD: Johns Hopkins University Press.
- Stein, Janice Gross, and Raymond Tanter. 1980. *Rational Decision Making: Israel's Security Choices, 1967*. Columbus, OH: Ohio State University Press.
- Steinberg, Blema. 1989. "Shame and Humiliation in the Cuban Missile Crisis: A Psychoanalytic Perspective." Paper presented at the annual meeting of the International Society for Political Psychology, Tel Aviv, Israel.
- Stoessinger, John. 1993. *Might of Nations*. 10th ed. New York: McGraw-Hill.
- Subrahmanyam, K. 1974. "An Indian Nuclear Force in the Eighties." In *The Superpowers in a Multinuclear World*, edited by G. Kemp and R.L. Pfaltzgraff, 127-40. Lexington, KY: D.C. Heath.
- 't Hart, Paul. 1994. "Coping with Groupthink in Government: Do We Have a Problem and If So, Should We and Can We Do Something about It?" Paper

- presented at the annual meeting of the International Studies Association, Washington, DC, 29 March – 1 April.
- Tanter, Raymond. 1979. "International Crisis Behavior: An Appraisal of the Literature." In *Studies in Crisis Behavior*, edited by Michael Brecher. New Brunswick, NJ: Transaction Books.
- Taylor, Stan A., and Theodore Ralston. 1991. "The Role of Intelligence in Crisis Management." In *Avoiding War: Problems of Crisis Management*, edited by Alexander George, 68–84. Boulder, CO: Westview Press.
- Tetlock, Philip E. 1987. "Testing Deterrence Theory: Some Conceptual and Methodological Issues." *Journal of Social Issues* 43: 85–91.
- Time*. 1985. "Who Has the Bomb: The Threat Is Spreading." 3 June, 36–52.
- Vandeputte, Edgard. 1995. Personal interview with the author, 29 June.
- Vasquez, John A. 1987. "The Steps to War: Toward a Scientific Explanation of Correlates of War Findings." *World Politics* 40: 108–45.
- 1993. *The War Puzzle*. Cambridge: Cambridge University Press.
- Verba, Sidney. 1969. "Assumptions of Rationality and Nonrationality in Models of the International System." In *International Politics and Foreign Policy*, edited by James N. Rosenau, 217–31. New York: Free Press.
- Wagner, R. Harrison. 1982. "Deterrence and Bargaining." *Journal of Conflict Resolution* 26: 329–58.
- Wallace, M. 1971. "Power, Status and International War." *Journal of Peace Research* 8: 22–35.
- Walters, R.E. 1974. *The Nuclear Trap: An Escape Route*. Harmondsworth, England: Penguin Books.
- Waltz, Kenneth. 1964. "The Stability of a Bipolar World." *Daedalus* 43: 881–909.
- 1979. *Theory of International Politics*. Reading, MA: Addison-Wesley.
- 1981. "The Spread of Nuclear Weapons: More May Be Better." *Adelphi Papers*. No. 171. The International Institute for Strategy Studies. London: Eastern Press.
- 1986. "Reflections on Theory of International Politics: A Response to My Critics." In *Neorealism and its Critics*, edited by Robert Keohane, 323–45. New York: Columbia University Press.
- 1993. "Toward Nuclear Peace." In *The Use of Force, Military Power, and International Politics*. 4th ed. Edited by Robert Art and Kenneth Waltz, 554–68. New York: University Press of America.
- Wayman, F.J. 1984. "Bipolarity and War: The Role of Capability Concentration and Alliance Patterns among Major Powers, 1816–1965." *Journal of Peace Research* 21: 61–78.
- 1985. "Bipolarity, Multipolarity, and the Threat of War." In *Polarity and War: The Changing Structure of International Conflict*, edited by Alan Ned Sabrosky, 115–44. Boulder, CO: Westview Press.
- Weede, E. 1981. "Preventing War by Nuclear Deterrence or by Détente." *Conflict Management and Peace Science* 6: 1–18.

- 1983. "Extended Deterrence by Superpower Alliance." *Journal of Conflict Resolution* 27: 231-53.
- Welch, D.A. 1989. "Crisis Decision Making Reconsidered." *Journal of Conflict Resolution* 33: 430-45.
- Weltman, John J. 1974. "On the Obsolescence of War." *International Studies Quarterly* 18(4): 413-26.
- Wilkenfeld, Jonathan, and Michael Brecher. 1979. "Superpower Crisis Management Behaviour." In *Foreign Policy: U.S./USSR*, edited by C.W. Kegley and P.J. McGowan, 185-212. London: Sage.
- 1988. "Superpower Crisis Management Behavior." In *Foreign Policy: U.S./USSR*, edited by C.W. Kegley and P.J. McGowan, 185-212. London: Sage.
- Williams, Phil. 1990. "Crisis Management in Europe: Old Mechanisms and New Problems." In *Avoiding War: Problems of Crisis Management*, edited by Alexander George, 500-18. Boulder, CO: Westview Press.
- Wittgenstein, Ludwig. 1968. *Philosophical Investigations*. 3d ed. Translated by G.E. Auscombe. New York: Macmillan.
- Wohlstetter, Albert. 1959. "The Delicate Balance of Terror." *Foreign Affairs* 37: 211-34.
- Wohlstetter, Albert, and Roberta Wohlstetter. 1971. "Controlling the Risks in Cuba." In *The Use of Force*, edited by Robert Art and Kenneth Waltz, 234-73. Boston: Little, Brown & Co.
- Wohlstetter, Roberta. 1962. *Pearl Harbour: Warning and Decision*. Stanford, CA: Stanford University Press.
- Wolfers, Arnold. 1959. "The Actors in International Politics." In *The Theory and Practice of International Relations*, edited by David McLellan, William Olson, and Fred Sonderman, 8-19. Englewood Cliffs, NJ: Prentice-Hall.
- Wright, Quincy. 1965. "The Escalation of International Conflict." *Journal of Conflict Resolution* 9: 434-49.
- Yermolenko, D.V. 1967. "Sociology and Problems of International Conflict." *International Affairs* 8: 14-19.
- Zacher, Mark. 1992. "The Decaying Pillars of the Westphalian Temple: Implications for International Order and Governance." In *Governance without Government: Order and Change in World Politics*, edited by James Rosenau and Ernst-Otto Czempiel, 58-101. Cambridge: Cambridge University Press.
- Zagare, Frank C. 1985. "Towards a Reformulation of the Theory of Mutual Deterrence." *International Studies Quarterly* 29: 155-69.
- Zandee, Dick. 1995. Interview with the author, 29 June.
- Zinnes, Dina. 1976. *Contemporary Research in International Relations*. New York: Free Press.
- 1980a. "Why War: Evidence on the Outbreak of International Conflict." In *Handbook of Conflict: Theory and Research*, edited by T.R. Gurr. New York: Macmillan.
- 1980b. "Three Puzzles in Search of a Researcher." *International Studies Quarterly* 24: 315-42.

Index

- Achen, Christopher, 5
Akashi, Yasushi, 133, 167n7
Axelrod, Robert, 58, 156n4
Azar, Edward, 156n6
- Berlin deadline crisis, 50
Betts, Richard, 29-30
Bosnian civil war, 120, 121, 124, 126, 127, 133; exchanges between Bosnian Serbs and U.S.-NATO forces, 127
Boutros-Ghali, Boutros, 133
Brams, Steve, 58-61, 64, 69, 81, 82, 83, 156n4,7
Brecher, Michael, 12, 48, 96-102, 106
Brodie, Bernard, 113
Bueno de Mesquita, Bruce, 31, 67, 95, 115, 149n5
- China, role in Korean war, 76
Cienfuegos Bay crisis, 76
coefficient of predictability and deterrence theory, 128-33
coercive diplomacy, success and failure of, 126
command, control, communication, and intelligence (C3I), 116, 117
Conflict and Peace Data Bank (COPDA), 86, 90, 91
conflict-cooperation transition, 89
conventional deterrence, xvi, 121-35
Corson, W.H., 81
Cox, Robert, 142
crisis decision making: and cognitive rigidity, 7, 10; groupthink effects on, 7, 8, 9; and motivational biases, 7, 10; and standard operating procedures, 90
crisis management and deterrence theory, impediments to progress, 6-17
crisis severity: and crisis provocation, 97, 100, 101; and geostrategic importance, 97, 98; and gravity of threat to superpower values, 97, 99-100; heterogeneity and, 97, 98; and number of issues, 97, 99; and perceived threat to superpower influence, 97, 100, 101; and linear and exponential threat ratings, 102
crisis stability, 114; in new nuclear states, 116
crisis violence, 104-6; centrality of, 102, 104; intensity of, 102, 104; and major response, 102, 103; and principal techniques of crisis management, 102, 103
Cuban missile crisis, 13-14, 25, 76, 81, 88, 92
cumulation: in international relations, xiv, xv, 86-93, 112, 140-2, 145nn1-2, 146n1; impediments to, 3-6; alternative approaches to, 17-18; in the social sciences, 142-3

- deductive and inductive approaches to theory, 3-6
- deterrence: rational theory of, 5-6, 85, 15, 158n1; and crisis management theory, 6-7, 14-17; success and failure of, 15, 19-20, 25, 26, 38, 45-6, 122-3, 150n12, 151n16; critics of, 16-17; and aggregate analysis, 20-32; classifying literature, 21; and credibility, 23, 24, 150n11; general versus immediate, 24, 25, 26, 27, 33; and compellence theory, 25, 121, 124; and expected utility theory, 28, 150n13; validity of aggregate testing, 32-5; and challenger/defender designations, 33, 126; and the role of nuclear weapons, 33; reliability of aggregate testing, 35-9; and the Syrian invasion of Lebanon, 39-46; and Israel's red-line conditions, 43-4; and intrinsic versus strategic interests, 45; and superpower rivalry, 52-5; coding cases of, 122-3, 150n14; alternative methods of testing, 122; and interpretations of historical events, 122; and enduring rivalries, 123-4; and selection bias, 123, 165n2, 166n4; requirements of, 125; coefficient of predictability for, 128-33; necessary and sufficient conditions for, 133
- dual-key system for UN-NATO peacekeeping operations, 133
- enduring rivalries, 123
- events data, 86-92
- extended deterrence, 22
- Feaver, Peter, 164n7
- Fleming, Paula, 9-10
- foreign policy crisis, definition of, 48
- Freeman, John R., 86-92
- Gaddis, John Lewis, 112-13
- game theory, 56, 66, 149n5; and nuclear deterrence, 54; and superpower rivalry, 56-72; chicken, 56, 60-1, 66, 155n1-3; prisoner's dilemma, 56, 60, 66, 85, 156n5; criticisms of, 66-7, 158n14; and nonmyopic equilibrium, 155n3; and stag hunt, 157-8n13
- Geller, Daniel, 32
- George, Alexander, 5, 14, 23-5, 64, 135, 151n16
- Geertz, Gary, 123, 124, 167n9
- Goldstein, S.J., 86-92, 157n9
- graduated and reciprocated initiatives in tensions reduction (GRIT), 54, 61-2, 68, 71, 79-80, 156n8; retaliation function for, 65; satisfaction and goal achievement with, 84, 85; GRIT-type strategies, 87-90
- GRIM, 54, 62-3, 65, 68, 71, 79, 80, 81, 157nn10-11; retaliation function for, 65; satisfaction and goal achievement with, 84, 85
- groupthink, 7, 8
- Gulf War, 12, 117, 118
- Hungary, Soviet invasion of, 25, 149n7
- Huth, Paul, 19-20, 25, 27-9, 30, 34-5, 124, 151n16
- intercontinental ballistic missiles (ICBMs), 115
- International Atomic Energy Agency (IAEA), 118
- International Crisis Behavior Project (ICB), 12, 45, 55, 74, 159n19, 160nn23-4
- inverted-U hypothesis, 14
- Israel: Six Day War, 38-9; Yom Kippur War, 39, 108-11; and the Syrian invasion of Lebanon, 41; and red-line conditions during Syrian invasion of Lebanon, 43-4
- James, Patrick, 96-100, 102, 136, 159n16, 167n11
- Janis, Irving, 7, 147n6
- Jervis, Robert, 5, 6, 8, 20, 45
- Kahn, Herman, 113
- Kennedy, John F., 14, 88, 156-7n8
- Kilgour, Marc, 58-61, 64, 69, 81, 82, 83, 156nn4,7
- Kissinger, Henry, 76, 77, 109, 110
- Korean War, 75-6
- Kugler, Jacek, 26-7, 94-5
- Kuhn, T.S., 145n2
- Lawson, Fred, 153n26
- Lebanon: Syrian invasion of, 39-46; civil war, 40, 152n20
- Lebow, Richard Ned, 5, 16-17, 38-9, 147n11
- Leng, Russell, 63, 81
- Lieberman, Eli, 123, 124
- long postwar peace, 163n2
- Lynch, Allen, 140-1

- margin of predictive error, 77-9
- Mearsheimer, John, xiv, 116, 143, 145n1
- meta-analysis, 148n13
- Morgenthau, Hans, 5
- Mueller, John, 62-3, 114, 157n11
- mutually assured destruction (MAD), 95, 143
- necessary and sufficient conditions for deterrence success, 133-4
- Nicholson, Michael, 11, 169n20, 170n21
- Nixon, Richard, 108, 109
- non-proliferation regimes, 117
- Non-Proliferation Treaty, 118
- North Atlantic Treaty Organization (NATO), 133
- nuclear crises: relative stability of, xv; as a test of deterrence theory, 52-5; severity and centrality of violence in, 55; provocation and retaliation levels during postwar, 72-7; and trigger-response levels, 74; and rational choice theory, 112-15
- nuclear deterrence: theory of, xv, xvi, 94, 96, 107, 154n2, 163n7; record of aggregate testing, 19-22; reliability of aggregate research, 35-7; case selection and coding, 47-52; and crisis management techniques, 51; and U.S.-Soviet rivalry, 94-108; logic of, 95; irrelevance of, 114; requirements of, 116
- nuclear proliferation: optimism and pessimism, 115-19; asymmetric, 118; increase in, 118; and international law, 119
- nuclear stability after the Cold War, 112-19
- Oneal, John, 11, 90, 164n3
- optimal threats (OT), 54, 58-61, 68, 77-8, 85; and the threat game, 60; retaliation function for, 65; and symmetric outcomes, 82; and satisfaction levels, 83
- Organski, A.F.K., 26-7
- Orme, John, 38, 39
- Osgood, Charles, 61-2, 68, 87, 88, 90, 161n30
- policy relevant theory, 135
- political psychology, 6, 7; versus rational choice, 7-14; critics of, 8-14; and crisis management theory, 14-17
- postmodernism, 137-42, 168n14; and linguistic relativism, 138; and cultural relativism, 139-41; and historical relativism, 142
- preventive diplomacy, 136
- protracted crises, 134; testing deterrence in, 122-35; and aggregate testing, 128; deriving consistent and inconsistent response sets for, 128-33; critics of, 134-5
- rational choice, 5, 6, 7, 10, 92; versus political psychology, 7-14; and crisis management theory, 14-17; versus non-rational models, 10, 11; and stress in foreign policy crises, 15; in nuclear crises, 112-15; critics of, 113
- Ray, James Lee, 88
- realist theories, 81, 112; enduring relevance of, 135-7; and state-centric assumptions, 136, 137, 167n12; postmodern critique of, 137-42; and unitary actor assumptions, 140
- reassurance theory, 16-17
- Riker, William, 31, 95
- Russett, Bruce, 19-20, 22-3, 25, 27-9, 34-5, 124, 151n16
- Sadat, Anwar, 39, 108, 156n8
- Sagan, Scott, 117, 119, 164n6, 165n11
- severity of nuclear threats, 96-102, 107
- Six Day War, 38-9, 48-9, 79
- Smith, T.C., 20-2, 148n2
- Smoke, Richard, 5, 23-35
- Snidal, Duncan, 5
- Soviet invasion of Afghanistan, 91, 92
- Soviet Union: fall of, xiii; nuclear capacity in 1949, 51
- status inconsistency theory, 4
- Stein, Janice Gross, 5, 16-17, 38-9, 147n11
- success/failure framework, 19
- Suez crisis, 162n2
- superpower rivalry: and postwar crises, 49, 160n23; and deterrence theory, 50-5; provocation and retaliation levels in, 65, 72-7, 91, 105, 113; and bargaining techniques, 73; and goal achievement, 81, 103-4; and symmetric ordering of alternative outcomes, 82; and integrative cumulation, 86-93;

- and short-term versus long-term cooperation, 88; and GRIT, EGRIT, and PGRIT strategies, 89; and reciprocity, 92; and the Yom Kippur War, 108-11; contemporary relevance of, 119-20
- Syrian invasion of Lebanon, 39-46, 152n25, 153-4n27; and Israel's deterrent threat, 42-5, 153n26
- system structure and stability, 4-5
- tit-for-tat (TFT), 54, 57-8, 65, 67, 77-8, 85; retaliation function for, 65
- Vance-Owen peace plan, 165n1
- Vasquez, John, 139
- Verba, Sydney, 10, 11
- Waltz, Kenneth, 5, 115, 168n12
- Watergate, 39
- Weede, E., 31-2
- Welch, David, 13
- Wilkenfeld, Jonathan, 48
- Wittgenstein, Ludwig, 138
- Wohlstetter, Albert, 116
- Wolfers, Arnold, 136
- World Events Interaction Survey (WEIS), 86, 87, 90, 91
- Yom Kippur War, 38-9, 50-1, 93, 94, 108-11, 151n16, 153n26; and DEFCON 3 alert status, 110; U.S. and Soviet involvement in, 108-11
- Zacher, Mark, 164n4
- Zagare, Frank, 155n3