1989

Methodological Themes and Variations

Philip E. Tetlock
University of Pennsylvania

Methodological Themes and Variations

Abstract

The future of civilization, perhaps of the human race, hinges on our ability to avoid nuclear war. This point has been made so repeatedly and occasionally so eloquently that it needs no amplification here (Dyson, 1984; Katz, 1982; Institute of Medicine, 1986; Sagan, 1983; Schell, 1980). The consensus, however, begins and ends on this point. There is wide disagreement on how likely nuclear war is, how such a war might occur, the forms it might take, and how it might best be prevented. Will nuclear war arise as the result of a "conflict spiral" between the nuclear superpowers—a self-reinforcing process driven by the tendency of each side to exaggerate the hostile intent of the other and to acquire ever more sophisticated weapons systems that increase the other side's sense of vulnerability and motivation to strike first in a crisis? Will nuclear war arise as a result of the failure of deterrence—a failure to convince the other side that one has both the political will and military capability to resist encroachments on "vital national interests"? Will nuclear war arise as a result of accident or miscommunication triggered by flaws in the command, control, and intelligence systems of the superpowers? Or has the casual role of the superpowers been overestimated? Will nuclear war arise from Third World conflicts of relatively remote relevance to U.S.-Soviet relations? And must nuclear war be an all-or-nothing proposition? Might limited nuclear wars periodically break out between particular powers or combinations of powers?

Disciplines

Applied Behavior Analysis | Defense and Security Studies | Emergency and Disaster Management | Energy Policy | Management Sciences and Quantitative Methods | Military and Veterans Studies | Peace and Conflict Studies | Policy History, Theory, and Methods | Science and Technology Policy

5

Methodological Themes and Variations

PHILIP E. TETLOCK

The future of civilization, perhaps of the human race, hinges on our ability to avoid nuclear war. This point has been made so repeatedly and occasionally so eloquently that it needs no amplification here (Dyson, 1984; Katz, 1982; Institute of Medicine, 1986; Sagan, 1983; Schell, 1980). The consensus, however, begins and ends on this point. There is wide disagreement on how likely nuclear war is, how such a war might occur, the forms it might take, and how it might best be prevented. Will nuclear war arise as the result of a "conflict spiral" between the nuclear superpowers—a self-reinforcing process driven by the tendency of each side to exaggerate the hostile intent of the other and to acquire ever more sophisticated weapons systems that increase the other side's sense of vulnerability and motivation to strike first in a crisis? Will nuclear war arise as a result of the failure of deterrence—a failure to convince the other side that one has both the political will and military capability to resist encroachments on "vital national interests"? Will nuclear war arise as a result of accident or miscommunication triggered by flaws in the command, control, and intelligence systems of the superpowers? Or has the causal role of the superpowers been overestimated? Will nuclear war arise from Third World conflicts of relatively remote relevance to U.S.-Soviet relations? And must nuclear war be an all-or-nothing proposition? Might limited nuclear wars periodically break out between particular powers or combinations of powers?

The contributions to *Behavior, Society, and Nuclear War* do not, of course, yield definitive answers to these questions. Given the multiplicity of potential determinants of the events in question and the imperfection of our knowledge of underlying causal processes, it is impossible to assign precise or even approximate probability estimates to alternative hypotheses concerning how nuclear war might occur (Allison et al., 1985). Accordingly, our goals here are both more modest and attainable. We draw on the behavioral and social sciences not to make precise predictions, but to underscore the enormous ambiguity and complexity of the threat of nuclear war, illustrate the types of research methods available for testing competing claims concerning how nuclear war might break out, highlight the limitations of simple single cause models of nuclear war, and clarify the variety of possible event sequences that could lead to such a war.

Intellectual honesty requires acknowledging that no one knows whether a nuclear war will occur or, if one occurs, what causes will have produced it or what forms the war will take. If nothing else, this acknowledgment serves as a sharp reminder to be wary of the pervasive human tendency to be overconfident in the correctness of one's forecasts and predictions (see Fischhoff, forthcoming). The acknowledgement should be taken, however, not as cause for epistemological despair, but as a useful starting point for analysis. There

has never been a nuclear war. The only instance in which nuclear weapons have actually been used in warfare were the atomic bombings of Hiroshima and Nagasaki in 1945. The United States—which possessed an overwhelming superiority in conventional weapons and a global monopoly on nuclear weapons—used the atomic bomb to force the surrender of an isolated and embattled Imperial Japan. We do not, fortunately, have a large data base of previous nuclear wars that we can probe and search for preconditions, correlates, or causes of the decision to use these extraordinarily destructive weapons. But if we lack in direct historical precedents, we risk being overwhelmed by the massive amount of indirect evidence at our disposal. There is no shortage of systematic high-quality research on international conflict (Levy, this volume). And there is no shortage of high quality research on the psychological, organizational, and political processes that make various forms of interpersonal and intergroup conflict more or less likely and more or less severe. If one is willing to grant that the same basic processes that shape these less apocalyptic forms of conflict also bear on the likelihood of nuclear war, then we have a sound logical basis for linking the behavioral and social sciences to the problems of both identifying plausible pathways to nuclear war and plausible preventive measures for avoiding such a war. To be sure, the linkages must be made cautiously, with sensitivity to the unique features of the nuclear predicament confronting us in the late twentieth century (for example, the tremendous accuracy and destructive power of the weapons, the rapidity with which decisions in nuclear crises might have to be made, and the danger of "losing control" over extremely complex technological systems). But the alternative to cautious inductive inference from "relevant" findings in the behavioral and social sciences also needs to be kept in mind. To place nuclear war in a category of its own—in which all we know about human behavior and society is peremptorily deemed irrelevant—is both difficult to justify and, in its most extreme form, a counsel of despair. How can one prepare to avoid an event that transcends existing theoretical and empirical knowledge? This chapter and, indeed, the entire series are premised on the assumption that although there will always be residual uncertainty about whether any given finding or generalization would hold up in a particular nuclear war scenario, we have learned a good deal about behavioral and societal processes that bear on the likelihood of nuclear war and, a central theme of this chapter, we have also learned a good deal about how to learn more about these processes.

This chapter has four specific purposes. The first objective is to provide an overview of the wide range of levels of analysis relevant to the problem of nuclear war. The overview will be sketchy, not exhaustive. The intent is to give a sense of the diverse disciplinary perspectives represented in this series. For analytic convenience, these perspectives have been divided into six (by no

means mutually exclusive) categories: cognitive and affective processes of individual decision makers, small group dynamics that emerge when decision makers interact, bureaucratic politics, problems of managing complex military-political organizations, domestic political processes, and the structure of the international system.

The second objective is to provide an equally broad overview of the research methods that investigators can employ to explore processes that could produce a nuclear war. We will consider the five most heavily used methods of generating such knowledge: laboratory experiments and simulations, mass surveys of public opinion, quantitative analyses of historical data, in-depth case studies, and formal mathematical models. The topic of research methods, it should, however, be stressed, cannot be approached in isolation from theoretical issues. One's theoretical assumptions are tightly coupled to one's methodological preferences. For example, investigators who believe that cognitive limits on rationality often play a key role in shaping national security decisions draw heavily from laboratory research on the judgment and choice processes of individual human beings; investigators who see such decisions as severely constrained by organizational and systemic processes draw heavily from research methods suited for studying more "macro" processes of that type (for example, single and comparative case studies and multivariate statistical studies of historical data).

The third objective is to highlight ways in which different research methods and theoretical perspectives complement and mutually enrich each other. It is tempting (especially given the insularity of citation practices in the behavioral and social sciences) to emphasize the difficulty of interdisciplinary communication. Researchers from different traditions often appear to speak different data-languages—languages so different that they cannot be translated into each other's language without egregious loss of meaning. This "incommensurability thesis" is too pessimistic. There is growing evidence of multimethod convergence on key causal propositions relevant to particular paths to nuclear war. There is also a strong logical and empirical case to be made for increased integration of theories in the behavioral and social sciences. "Macro" theories can benefit by building on more realistic assumptions concerning the nature of the human decision maker (Simon, 1985); "micro" theories can benefit by taking into detailed account the normative and institutional constraints on decision making and the nature of the problems with which decision makers must cope (Pfeffer, 1985).

The fourth and final objective is to explore the problem of "policy relevance." Behavioral and social scientists are typically concerned with identifying empirical generalizations and theoretical propositions that hold up over large classes of observations (whether the units of observation be individuals,

institutions, nation-states, or even international systems). Policymakers, on the other hand, want to know how to deal with a specific adversary on a specific issue at a specific time. The chapter concludes by examining the complex conceptual and methodological obstacles that arise in applying behavioral and social science knowledge to specific policy controversies.

Levels of Analysis

The study of international conflict can be likened to looking through a microscope at different levels of magnification. At the highest levels of magnification, the focus is on the cognitive, emotional, and even psychophysiological responses of the individual decision maker (Axelrod, 1976; Etheredge, 1978; Holsti, 1976; Lebow, 1981; Jervis, 1976; Walker, 1977; Wiegele, 1979). As one reduces the magnification, one loses the ability to draw fine-grained inferences about individual states of mind but gains the ability to place events in successively broader systems contexts. Thus, one no longer has access to data that allow one to test detailed hypotheses concerning the content of individual belief systems, the nature of the decision rules used to rule out options, or the motives that underlie policy preferences. One gains, however, the ability to draw inferences concerning the impact of the surrounding context on the decision process. The key advisors to the central decision maker, and the interpersonal and group dynamics among them, initially come into focus (Janis, 1982). These small group processes affect both the options considered and the procedures used for assessing those options. Next, the key agencies of government—with individual policymakers now appearing as role representatives of powerful bureaucratic constituencies—come into focus (Allison, 1971; Halperin, 1974; Art, forthcoming). Next, the domestic political and economic environment surrounding the central government looms into view (Russett, this volume). In the case of the United States, we must now take into account Congress, public opinion, defense contractors, the news media, mass political movements, and the interactions among them. Finally, we are confronted by the international system: the complex web of economic, political, and cultural entanglements that define one nation's relationship to another. We must also consider the balance of power and the many variables impinging on that balance: the conventional and nuclear forces of one's own and of other nations, patterns of interaction among nations (trade, formation and dissolution of alliances), and the economic underpinnings of military strength (growth rates, technological development, access to critical natural resources).

It is not surprising that communication across levels of analysis is both

difficult and rare. What excites the attention of investigators working at one level of analysis may well be invisible to investigators working at other levels of analysis. One can study foreign-policy decision making at a purely cognitive level of analysis without ever referring to research on group dynamics, role theory, bureaucratic politics, special interests, public opinion, trading patterns, or the balance of power. The substantive content of the decisions would, in a fundamental sense, be irrelevant. The "cognitivist" would be concerned solely with the types of decision rules employed, the strategies used to cope with uncertainty and trade-offs, the degree of openness to new evidence, etc. Conversely, one can study patterns of interaction among nations—alliance formation and dissolution, the waxing and waning of arms races—without ever referring to cognitive research on belief systems or judgmental heuristics. Researchers at these higher levels of analysis often feel comfortable working with very simple assumptions concerning the human decision makers involved. It suffices to posit a capability for "rational" thought, a concern for power and, perhaps, an attitude toward risk.

The levels-of-analysis problem is a familiar one to international relations researchers, and there is not much to be gained from an extended examination of this much debated and essentially unresolved controversy (Greenstein, 1975; Hoffman, 1960; Holsti, 1976; Jervis, 1976; Kelman and Bloom, 1973; Rosenan, 1966; Singer, 1972; Verba, 1961; Waltz, 1959). There is still little agreement on the relative contributions that different levels of analysis can make to the explanation of international conflict. There are, however, at least signs of growing tolerance among advocates of competing schools of thought. Many scholars now agree that the importance of any given level of analysis is not a constant but is likely to vary as a function of the configurations of variables at other levels and the types of questions we want answered (Holsti, 1976; Jervis, 1976; Levy, this volume). For instance, systemic variables may shape the major challenges confronting a nation's foreign policy and the overall direction its policy may take, but not its exact responses in a specific situation. Thus, economic and geopolitical pressures might have made a major European war in the early twentieth century very likely (Choucri and North, 1975), but a psychological analysis of crisis decision making and of the personalities involved may be needed to explain why World War I broke out in August 1914 (Holsti, 1972). Or, to take another example, systemic variables may exert a powerful constraining influence over the foreign policies of individual nations, but that constraining influence may not be sufficient when variables at lower levels of analysis take on extreme values (for example, Hitler's Germany of 1939, or Khomeini's Iran of the 1980s).[1]

A great deal of the research reviewed in *Behavior, Society, and Nuclear War* points to the need for complex or contingent generalizations of this sort

(George, 1980). Different levels of analysis, in this view, do not represent mutually exclusive ways of looking at the world. The intensity of international conflict and the diverse forms it takes are typically the product of interactions among variables at many levels of analysis. In Allison's (1971) metaphor, levels of analysis are best thought of not as rival theories but as "beacons" that sensitize investigators to different bodies of data, research methods, and potential explanations. Psychological explanations lead one to expect patterns of individual and small group behavior among policy elites that one simply would not have expected if one restricted theorizing to a realpolitik, balance-of-power framework. If one assumes that foreign policy is the product of unitary, rational, power-maximizing actors, there is no reason for expecting decision makers to fall prey to serious misperceptions (Holsti, this volume; Jervis, 1976; Stein, forthcoming) or to engage in self-defeating patterns of behavior in small groups (Janis, 1982) or to advocate narrowly defined bureaucratic interests at the expense of broader national policy objectives (Halperin, 1974). These latter findings can be quite readily explained, however, if one grants the possibility that certain basic laws of psychosocial functioning hold up in very different spheres of life (for example, in laboratory experiments and cabinet rooms). Conversely, systemic theories lead one to expect regularities in international conduct that individual-level theorists would probably have never anticipated. For instance, Bueno de Mesquita (1985) has shown that it is possible to achieve impressively accurate predictions of the initiation of war based on purely systemic indicators (balance-of-power measures) with the assistance of only the most rudimentary psychological assumptions. Modelski (1987) and others have demonstrated intriguing regularities in the outbreak of major wars over the last five centuries, tracing these patterns to fundamental political-economic processes that make the decline of hegemonic powers inevitable.

This "beacon" interpretation suggests that the best test of a level of analysis is the power of theories formulated at that level to stimulate the discovery of important findings that otherwise would have remained undiscovered. The central research challenge is not to debunk competing levels of analysis but to improve the quality of theoretical and empirical work at one's chosen level. Efforts to achieve decisive tests between levels of analysis may be superficially appealing (create the appearance of cumulative hypothetico-deductive science) but actually may be extremely misleading. Confrontations between levels of analysis are often premature. They presume methodological refinement that just does not exist. Demonstrations that variables from one level of analysis "outpredict" variables from another level may tell us little about the explanatory superiority of that level and a lot about the relative sophistication of theory development within the two levels at a given time or the relative

sophistication of techniques for operationalizing variables within the two levels at a given time. It is unwise to draw strong conclusions about the long-term explanatory potential of levels of analysis from such studies. The risks of underestimating the explanatory potential of the early losers and of overestimating the potential of the early winners are too great. A more prudent strategy is to encourage investigators to put their own "theoretical houses" in order before trying to annex new explanatory territory. As we will see in *Behavior, Society, and Nuclear War*, investigators working at the same level of analysis frequently find themselves in sharp disagreement over what variables are most important, how variables interact, and how these variables should be assessed.[2]

Next, I briefly sketch the types of processes studied at each of six different levels of analysis and the types of methods used to study these processes.

Cognitive and Affective Processes of Individual Decision Makers

Many behavioral scientists take it to be self-evident that microlevel explanations—those that invoke the beliefs, values, perceptions, and feelings of individual decision makers—can contribute to our understanding of how to prevent nuclear war. What, indeed, could be more obvious? Policymakers are human beings. The behavioral sciences have been at least partly successful in identifying lawful regularities in human behavior. It seems to follow that the behavioral sciences are well positioned to shed light on both the causes of war in general and, by implication, the potential causes of nuclear war.

This reductionist argument probably proves too much. For, by the same logic, behavioral science is reducible to biochemistry (of what, after all, do people consist?) which, in turn, is reducible to subatomic physics (of what ultimately do complex protein chains consist?). It is not enough to assert a reductionistic claim; one must systematically document exactly how concepts and research methods from the putatively fundamental discipline help to clarify problems that arise at the next higher level of analysis.

Analysts of political decision making have attempted to do exactly that (Axelrod, 1976; George, 1980; Jervis, 1976). They have built a strong case that important similarities exist between decision making in foreign policy and in other spheres of life. Thus, they have noted the extraordinary difficulty of identifying a best or utility-maximizing solution to most foreign policy problems (Steinbruner, 1974). Policymakers must deal with incomplete and unreliable information on the capabilities and intentions of other states (sometimes even of their own states). The range of response options confronting them are indeterminate. The probable consequences of each option are shrouded in uncertainty. Policymakers must compare options on many con-

flicting, seemingly incommensurable, value dimensions (for example, the impact of options on economic interests, international prestige, domestic popularity, human rights, and even lives). Finally, to compound the difficulty of the task, policymakers must sometimes work under intense stress and time pressure (Holsti, this volume; Janis and Mann, 1977; Lebow, 1981; Stein, forthcoming; Suedfeld and Tetlock, 1977).

Advocates of information processing explanations argue that policymakers frequently resort to simplifying strategies to deal with the complexity, ambiguity, and painful trade-offs inherent in foreign policy problems (Fischhoff, forthcoming). These simplification strategies can take many forms: reliance on simple historical analogues or precedents in interpreting new situations (Neustadt and May, 1986), reluctance to modify preconceptions in response to challenging evidence (George, 1980; Jervis, 1976), perceiving situations in ways designed to minimize evaluative inconsistency and value trade-offs (Axelrod, 1976), and dependence on simple, easy-to-execute heuristics in assessing the likely future behavior of other states (Jervis, 1976). It is also argued that cognitive economy and efficiency frequently have a steep price: susceptibility to error. Reliance on simple historical analogies raises the risk of overlooking important differences between one's preferred precedent and the current problem (for example, the Vietnam War differs in many respects from the diverse contemporary conflicts to which it has frequently been compared: El Salvador, Nicaragua, Afghanistan, Lebanon, Angola, Ethiopia, and Cambodia). Reluctance to modify preconceptions raises the risk of clinging to incorrect assessments of situations in the face of unexpected developments (for example, U.S. analysts who firmly believed in the monolithic nature of the Communist movement were slow to recognize the strategic significance of the Sino-Soviet dispute). Intolerance of evaluative inconsistency raises the risk of failing to recognize flaws in policies one supports and virtues in policies one rejects. Dependence on simple attributional heuristics raises the risk of drawing too sweeping and confident inferences concerning how others are likely to act in the future.

Balancing the possible benefits of cognitive efficiency against the possible increases in error is a perplexing normative problem with no widely acknowledged solution. By contrast, documenting cognitive constraints on judgment and choice processes is a relatively straightforward empirical problem with a quite widely acknowledged methodological solution. The solution takes the following form: (1) demonstrating—largely through experimental research—that people rely on certain cognitive strategies to cope with complexity and uncertainty; and (2) demonstrating—largely through historical case studies and content analyses of policy deliberations—that foreign policy decision makers appear to rely on similar sorts of strategies. In short, one looks for multimethod convergence.

This research strategy is especially appealing because it builds on the complementary strengths of experimental research (the ability to test detailed process models of individual behavior and to eliminate alternative causal hypotheses) and of naturalistic research methods (much more immediate relevance to events and issues bearing on the likelihood of nuclear war). One is justified in holding greater confidence in generalizations concerning judgment and choice that successfully pass two so very different methodological screening tests. And there is considerable evidence that the generalizations mentioned earlier hold up quite well indeed (George, 1980; Jervis, 1976; Tetlock and McGuire, 1986).

Although multimethod convergence of this sort is encouraging, caution is still appropriate for two important reasons. For one thing, it is possible to have spurious convergence. Naturalistic researchers might wrongly conclude that foreign policy behavior that only superficially resembles a laboratory analogue is the product of the same underlying cognitive or affective process. Thus, policymakers may appear to rely on simple rules of thumb in drawing lessons from history, but they may actually be working with a far more subtle and sophisticated grasp of the situation. Policymakers may be using simple historical arguments (such as "no more Munichs" or "no more Vietnams") to rally support from wavering political constituencies and to preempt potential criticism from either the left or right. In a similar vein, policymakers may not actually be unaware of value trade-offs or of contradictory evidence but may find it politically useful to refuse to acknowledge them. Distinguishing between genuine and spurious multimethod convergence (or, in this case, between perceptual-cognitive and political impression management explanations) is often a very tricky judgment call that requires detailed knowledge of specific historical episodes.

It is also a judgment call that many contributors to our series must implicitly or explicitly make. Holsti (this volume), for example, needs to make such judgments in assessing whether crisis-induced stress really impairs the judgment of foreign-policymakers or whether foreign-policymakers are craftily trying to influence the calculations of other national leaders by persuading them that such impairment has occurred (see Schelling, 1966, on the rationality of occasionally appearing irrational). Stein (forthcoming) needs to make similar sorts of judgments in assessing whether the leaders of states that challenge deterrence are allowing motives and wishes to distort their perception of the choices available to them or whether they are cleverly trying to intimidate the status quo power by persuading it that they have no choice (given domestic pressures) but to persevere with confrontational policies. We confront variants of the same analytic dilemma in the Fischhoff (forthcoming) chapter. How can we assess, for example, whether decision makers in nuclear command, control, and communications systems are overconfident in their

ability to avoid Type I errors (falsely conclude that an attack is occurring) and Type II errors (falsely conclude that an attack is not occurring) or whether these decision makers are self-consciously promoting a necessary social-political fiction? In short, it is no simple matter to determine whether one has discovered true multimethod convergence.

There is also a second major reason for caution. Multimethod divergence sometimes occurs. For most laboratory-based generalizations, it is fairly easy to identify numerous exceptions and qualifications in the "real world." Decision makers sometimes draw flexible, multidimensional lessons from history (Neustadt and May, 1986), confront trade-offs even in highly stressful situations (Maoz, 1981), and display a willingness to change their minds in response to new evidence (Tetlock, 1985b). Multimethod divergence of this sort does not, of course, mean that one set of findings must be "right" and the other "wrong." When radically different research methods are being compared, an enormous number of possible explanations exist for divergence. How people think may depend on a variety of boundary conditions: individual differences in cultural background, intellectual capacity, and cognitive and interpersonal style, and situational variables such as the nature of the decision-making task, task importance, small group processes, and role and accountability relationships. Each class of variable—by itself or in combination with others—may help to explain inconsistencies in the pattern of evidence.

Small Group Processes

Decision makers do not operate in a social or institutional vacuum. Most important national security decisions appear to be collective products—the result of often intensive interactions among small groups of decision makers, each of whom in turn represents major bureaucratic or political constituencies. The norms and operating procedures of these small groups thus become potentially important determinants of policy outcome.

Evidence at hand suggests that these small group processes can interact in many ways with individual dispositions to influence the decisions that policymakers reach. Janis (1982), for example, reviews a considerable body of both historical and experimental research that suggests that, under certain conditions (directive leadership, cohesive group, high external threat, etc.), group norms will emerge that exacerbate already dangerous trends in individual judgment. Far from checking bias and error in each other, policymakers in these "groupthink" situations behave in ways that encourage overconfidence, self-righteousness, cognitive rigidity, and excessive optimism and that discourage dissent and the expression of unpopular doubts or opinions. The

result, according to Janis, is the undertaking of ill-thought-out foreign policy projects that frequently lead to disastrous consequences (for example, the pursuit of the defeated North Korean army deep into the Korean peninsula in 1950 and the provoking of Communist Chinese intervention or the abortive Bay of Pigs invasion of Castro's Cuba in 1961).

Small group processes do not always, however, make matters worse. Under other conditions (external accountability checks, nondirective leadership, institutional mechanisms for ensuring the representation of different points of view), group norms can facilitate complex, open-minded analyses of policy options (Janis, 1982; George, 1980: Chap. 11) and confer at least some protection against well-documented judgmental biases such as overconfidence, intolerance of evaluative inconsistency, and the tendency for first impressions to persevere in the face of later contradictory evidence (Tetlock, 1985b; Tetlock and Kim, 1987). Janis (1982) cites the development of the Marshall Plan and the handling of the Cuban missile crisis as exemplary models of how group processes can improve the quality of decision making (see also Holsti, this volume).

Investigators who seek to apply knowledge of small group dynamics to the problem of preventing nuclear war generally adopt the same methodological strategy as those who seek to apply knowledge of individual-level processes. On the one hand, there is a body of experimental research on topics such as conformity, coalition formation, leadership styles and effectiveness, minority influence, and the group-induced attitude polarization effect. On the other hand, there is a body of historical and archival research that probes the relevance of this work on "basic processes" for understanding the making of critical national security decisions (Burnstein and Berbaum, 1983; George, 1980; Janis, 1982; Tetlock, 1979). To the degree these very different strategies of inquiry point to compatible conclusions, our confidence in those conclusions should be reinforced. Nonetheless, all of the cautionary caveats appended to the multimethod convergence argument in the previous section apply here.

Decision Makers as Representatives of Complex Bureaucratic Constituencies

National security decisions can be viewed as outputs of complex military and political organizations—organizations such as the Department of State, Department of Defense, Central Intelligence Agency, National Security Council, and so forth. It is typically assumed that attributes of individual decision makers cease—at this level of analysis—to be terribly important; what matters are the complex bureaucratic and political perspectives and

interests that these decision makers are expected to represent. Decision makers become role incumbents; as the saying goes, "where they stand depends on where they sit."

Research on organizational behavior and bureaucratic politics abounds with examples of policymakers shifting supposedly firm convictions on assuming new posts or after shifts in the prevailing political atmosphere (Halperin, 1974; Kaplan, 1982). Such shifts should not be too surprising. There are often overwhelming normative pressures on decision makers to adopt attitudes consistent with both their own career and institutional interests (pressures, for example, to promote institutional control over policy and to justify appeals for increased budget and staff). Moreover, the readiness with which many high-level decision makers strategically shift their views is very consistent with what is known of the readiness of ordinary people to modify their opinions and conduct to the social demands of the moment. Large experimental literatures exist on ingratiation, self-presentation, and strategic attitude shifts—all of which underscore the plasticity of beliefs, attitudes, values, and feelings that one might otherwise suppose to be central to people's self-concepts (Schlenker, 1985).[3]

From the perspective of understanding determinants of policy preferences (which, in turn, may bear on the likelihood of nuclear war), the appropriate focus is on the institutional coalitions or systems that key decision makers represent. There is, for example, much work on the ways in which the institutional interests of the military affect the processes of weapons development, weapons procurement, and the setting of budgetary priorities (Melman, 1970; Sarkesian, 1972; Yarmolinsky, 1970). In recent years, there seems to have been a general tendency to favor offensive systems and, within that category, to favor systems that will expand the budgetary and manpower requirements and the political influence of one's own branch of the services. From the standpoint of students of organizational behavior, the tendency of national security bureaucracies to behave this way is no more surprising than would be the tendency of individual decision makers to rely on simple heuristics to cognitive psychologists. A basic principle of organizational functioning—the tendency to defend and, when feasible, expand claims on resources through socially acceptable justifications—has once again been validated (Pfeffer, 1985).

Documenting cause and effect relationships in complex bureaucracies is, of course, difficult. The available data are not impressive. One must rely largely on detailed interviews with former policymakers—whose memories and motives may be suspect—and on careful checks of archival records to cross-validate conclusions drawn from interviews. Nonetheless, the data can be suggestive. Correlations can be documented—through qualitative or quantitative, cross-sectional or time-series methods—between institutional self-

interest and institutional policy advocacy. For instance, if the positions taken by the Air Force and Navy on the military value (even morality) of countervalue nuclear strikes depends on the technological sophistication of the weapons systems available to those services at different times—namely, the accuracy of land-based international ballistic missiles (ICBMs) and the availability and, later, the accuracy of submarine-launched ballistic missiles (SLBMs)—that finding counts as suggestive evidence that institutional interests played a part in shaping critical national security estimates and priorities.

Managing Complex Military-Political Systems

High-level decision makers sit atop organizational systems of enormous complexity—systems whose intricacies it is unrealistic to suppose they can fully master given the multifarious demands on their time and their often brief tenures in key positions. As a result, decision makers may be unaware of how quickly they can lose control over the chain of causal events leading up to war. Once decision makers commit themselves to a particular line of action (for example, partial mobilization), that commitment may trigger a series of preprogrammed action-reaction cycles that make conflict all but inevitable. Historical analysts have made this argument repeatedly in the case of World War I (emphasizing the rigidity of mobilization schedules and railroad timetables); contemporary strategic analysts have made analogous arguments concerning current procedures for the command and control of nuclear forces (Blair, 1987; Bracken, 1983; Steinbruner, 1985, 1987). High-level decision makers can be easily overwhelmed by the technical complexity of the weapons systems, the tendency to decentralize authority over nuclear weapons in crises, the problem of "alert instability," and uncertainty concerning the "rules of engagement" for military forces.

Understanding how nuclear war might occur requires, from this perspective, absorbing a staggering amount of detailed technical information on how the military forces of the key nation-states actually function and of the forms that military-political coordination takes within those states. Research methods from the social sciences can, however, help us to cope with this potential information overload. Rigorous game-theoretic modeling can highlight how the adoption of particular policy proposals, such as highly accurate counterforce weapon systems (e.g., the MX) or various types of ballistic missile defenses, affect the incentives for each to launch preemptive strikes at various points in the escalation sequence (Brams, 1985; Shubik, forthcoming). Analyses of this sort can sensitize the policymaking community to how ostensibly purely "defensive" measures by one's own side can dramatically affect the relative attractiveness of options open to the other side.

Historical case studies of crisis decision making can also play a valuable

role. It is useful, for example, to be reminded of how standard military operating procedures have led, in the past, to incidents that could have triggered dramatic escalations of crises (for example, the need for careful oversight of the implementation of the naval blockade of Cuba to prevent an early clash between U.S. and Soviet forces during the missile crisis). Complex military operations send out a variety of both advertent and inadvertent signals to prospective adversaries—signals that can be easily misinterpreted (Thies, 1984). The leaders of other states may see a message being sent where none is intended (for example, a pause in bombing during the Vietnam War due to poor weather might be seen as a conciliatory gesture), or they may seriously misinterpret a message that is actually being sent (for example, a partial activation of military units might be seen as a prelude to a full-scale attack, even though it is intended only to indicate firmness of resolve).

The Nature of the Nation-State

The notion that some nations are intrinsically more prone to war than others is a long-standing one in the study of international politics (Waltz, 1959). Hypotheses on this issue take many contradictory forms. It has been suggested, for example, that capitalist states have special economic needs for expansion or for disposing of surplus capital (needs that predispose them to war). And it has been suggested that capitalism in conjunction with free international trade is actually the best guarantee of peace. It has been proposed that democratic political institutions encourage peace, and it has been proposed that democracies are especially vulnerable to volatile swings in public opinion that make war more likely. Finally, it has been argued—by various investigators at various times—that internal political conflict increases, has no effect on, and decreases the likelihood of war. There is, in brief, no shortage of speculation on the subject (see Levy, this volume).

Research generally takes one of three distinct forms at this level of analysis. One common form is the case study. Do national leaders in specific historical situations feel compelled by domestic political imperatives to pursue policies that are likely to result in war? In her chapter in the second volume of this series Stein reviews considerable evidence that this does indeed sometimes happen.[4] A second common form of research is the multivariate correlational study that probes for relationships between attributes of large numbers of nation-states (their form of government, level of economic development and type of economic organization, intensity of internal conflict, governmental stability, etc.) and the involvement of these nations in war. The third genre of research is the survey. Using relatively standardized sampling and interviewing techniques, the goal is to understand fluctuations in public opinion on foreign policy issues over time (for example, does out-group threat increase sup-

port for in-group symbols of authority such as the presidency?) and the relationship between public opinion and foreign policy (see Russett, this volume).

It is worth noting that an interesting example of both multimethod convergence and divergence emerges at this level of analysis. On the one hand, multivariate correlational studies of war lead us to conclude that the relationships between internal conflict within nations and involvement in external conflict are extremely weak, even nonexistent. On the other hand, experimental research, survey research, and historical case study research point to a different conclusion. Experimental work suggests that in-group cohesiveness increases as a function of external threat under a fairly broad range of circumstances (Brewer and Kramer, 1985; Coser, 1956). Survey research documents a compatible finding—the tendency for support for presidents to increase immediately after the use of force abroad (Mueller, 1973; Russett, this volume). And historical case study research suggests that high-level decision makers have a good intuitive grasp of this rather robust social science generalization and are willing on occasion to risk and even to seek external conflict in order to promote internal political cohesion and, by implication, their own political positions (Rosecrance, 1963; Lebow, 1981; Stein, forthcoming).

How should the apparent inconsistency be explained? There are several possibilities. One is that the inconsistency is illusory. The temptation to use external conflict to promote internal cohesion is real, but national leaders act on the temptation only when a variety of elaborate preconditions are met (for example, alternative, less risky methods of reducing internal conflict do not exist, a plausible pretext does exist for picking a quarrel with another power, the costs of quarreling are not prohibitive, and the internal conflict has not reached a point where it impairs the nation's capability and resolve to engage in external conflict.). As a result, internal conflict by itself is not a powerful predictor of external conflict. There are just too many—and too difficult to operationalize—moderator variables of the relationship between internal and external conflict.

Another possibility is that the multivariate correlational studies suffered from serious flaws that undermined their power to detect true relationships between internal and external conflict. Most studies, as Levy (this volume) points out, were based on a narrow and remarkably peaceful period in international politics (1955–1960). Since it is obviously unwise to draw sweeping conclusions from so brief and unrepresentative a slice of time, final verdict on this issue must await better cross-sectional and time-series evidence.

The Structure of the International System

Theories that are derived from this most "macro" level of analysis pay little or no attention to the psychological, institutional, and political processes that

we have considered at the lower levels. Whether a nation goes to war is studied without reference to the psychological processes of individual decision makers, small group dynamics among those decision makers, intragovernmental factionalism, the difficulties of managing enormously complex military-political systems, and domestic political conflicts within the state. The nation-state is treated as an undifferentiated atom—a unitary rational actor that attempts to maximize its influence within the constraints posed by the international system. Key explanatory variables from this viewpoint are system-level variables: the bipolarity versus multipolarity of the distribution of power (Waltz, 1979), the intensity of competition for raw materials and markets (Choucri and North, 1975), threats to the current balance of power (when one power appears close to achieving hegemony, a military coalition of other powers will often emerge to resist it), the relative rates of change in the military-economic power of key nation-states (Kennedy, 1987; Organski and Kugler, 1980), and the existence of alliances and the cohesiveness and stability of these alliances. Investigators can test systemic hypotheses using qualitative or quantitative data analytic techniques. Thus, research can take the form of detailed case studies of shifting international coalitions prior to a specific war or multivariate statistical studies of the power of systemic variables to predict the outbreak of war among large numbers of nations or over long periods of time (Choucri and North, 1975; Bueno de Mesquita, 1981, 1985; Organski and Kugler, 1980). Such studies have shed light on the conditions under which systemic processes such as competition for scarce resources or shifting balances of power lead to war—although theorists working at this level of analysis still disagree sharply not only over the magnitudes of particular effects but sometimes even over the directions of those effects (see Levy, this volume).

Other research possibilities also exist. For instance, it is possible—because of the strong assumptions some systemic theorists make concerning the unitary rationality of foreign policies—to explore the logical implications of systemic formulations through game-theoretic techniques. Granting that interdependency in the international environment takes a certain form (obviously a difficult judgment call), it is important to know whether there is a Nash equilibrium point in the payoff matrix, a set of policy postures from which no player would have a strong incentive to depart because of the substantial risk of shifting strategy (for example, the defect/defect option in the prisoner's dilemma game; see Shubik, forthcoming). It is also important to know whether the equilibrium point is Pareto-inferior (there are sets of policy postures that would leave at least some players better off at no one's expense *if* everyone acted in the prescribed way) or Pareto-superior (there are no sets of policy postures that yield better outcomes for *both* players). Many regularities observed at the systemic level—arms races, the formation of alliances, preemp-

tive wars against would-be hegemonic powers—can be viewed as the result of a tendency for rational national actors to plan for the worst (what game theorists term the minimax strategy of trying to maximize the "goodness" of one's worst possible outcomes). Pursuit of a minimax strategy may be rational in the sense of protecting oneself from exploitation, but often only at the price of locking each side into a Pareto-inferior set of outcomes (for example, never-ending arms races and geopolitical competition) and of precluding the identification of Pareto-superior sets of outcomes (for example, verifiable arms control or troop reduction agreements).

Research Methods

Research relevant to the problem of preventing nuclear war covers an enormous range of methodologies. As already noted, one's choice of research method hinges, in part, on one's level of analysis which, in turn, is often related to the nature of the outcome (dependent variable) one is trying to explain. It is possible to conduct highly controlled laboratory tests of certain causal hypotheses (for example, those pertaining to the impact of information load on choice strategies or the impact of pursuing a fair-but-firm reciprocity strategy on behavior in a mixed-motive game), but not of others (for example, one cannot experimentally manipulate the bipolarity-multipolarity of an international system or the degree of economic interdependence within an alliance).

In addition, one's choice of research method depends on one's epistemological preferences. Some investigators advocate extending the hypothetico-deductive or covering-law model of explanation to the study of international conflict. According to this model, an event has been explained only when one has shown that the existence of the event could have been inferred—either deductively or with a high probability—"by applying certain laws of universal or statistical form to specified antecedent circumstances" (Hempel, 1965:229). Advocates of this explanatory approach generally put their methodological faith in controlled laboratory experimentation (what better way is there to test causal hypotheses of universal or statistical form?) and in quantitative studies of the historical record (what better way is there to determine whether predicted lawful regularities actually hold up in the real world?). A pithy expression of this epistemological view is to be found in S.S. Stevens' famous critique of qualitative research: "when description gives way to measurement, calculation replaces debate" (in Kaplan, 1964:174). The key to cumulative scientific progress lies in identifying and quantifying lawful regularities in the phenomena under study.

Other investigators are markedly less enthusiastic. Little can be learned, in their view, from attempting to extend the covering-law model from domains where it has served us well (in many areas of the physical and biological sciences) to domains where it is inappropriate—inappropriate because of the complexity, uniqueness, and perhaps even indeterminacy of the underlying processes at work (Almond and Genco, 1977; Cronbach, 1975; Gallie, 1968; Gergen, 1978). Variable-centered, nomothetic research cannot begin to capture the multiplicity of motives activated in particular foreign policy problems, the confusion and uncertainty that frequently infuse the decision process, the norms and implicit understandings that exist among key actors, the institutional constraints on decision makers, and so forth. If we want to understand how nuclear war might break out, we need to understand in rich, idiosyncratic detail how policymakers live and work. The goal should be the "thick description" of specific events (Geertz, 1973) or the development of "coherent whole explanations" (Walsh, 1967) that trace the connections among events and then reveal superordinate themes that give meaning and context to those events.

It is, to be sure, misleading to divide the epistemological universe into two hostile, hopelessly incompatible camps. It is possible to see value in both quantitative, variable-centered research and in qualitative, idiographic research; it is even possible that the two lines of inquiry will occasionally lead to similar conclusions. Patterns or themes that emerge from in-depth case studies are sometimes highly consistent with lawlike generalizations that emerge from experiments or multivariate field research (for examples, see Druckman and Hopmann, this volume; Holsti, this volume; Stein, forthcoming). Although the epistemological packaging is certainly different, there is presumably a common underlying reality to which practitioners of these different approaches are responsive.

Most investigators—including the contributors to this series—do not appear to have made absolute or dogmatic commitments to a particular method of generating knowledge. Most seem to be pragmatists who are willing to shift methodological strategies depending on the nature of the problem under investigation and the available data. There is also a growing recognition that the distinctions among alternative research methods are not as clear-cut as often implied. It is possible to test hypotheses derived from covering-law explanations through comparative case studies. (For discussion of the method of focused comparison, see George [1979] and George and McKeown [1985]; for use of that method to identify problems that arise in applying abstract deterrence theory to specific historical cases, see George and Smoke [1979].) And there is no reason why quantitative, variable-centered researchers cannot be more sensitive to potential boundary conditions on the applicability of their

hypotheses or to the problems that arise in operationalizing abstract constructs in concrete situations. (See Druckman and Hopmann [this volume] on the need to strike a balance in content analyzing negotiations between sensitivity to the uniqueness of the particular case and to the need for theoretical generality.)

Next I consider the range of research methods that can be deployed to increase our understanding of how nuclear war might occur. These methods differ from each other in many ways. They vary in the degree to which they permit confident causal inferences (well-controlled, laboratory experiments conferring the greatest inferential power, single historical case studies generally the least, with multivariate field studies somewhere in between), in their immediate or obvious relevance to the problem of avoiding nuclear war (historical case studies of U.S.-Soviet relations possessing the most relevance, laboratory experiments of college students perhaps the least), in their reliance on qualitative versus quantitative forms of argument (laboratory researchers and game theorists tending to be the most mathematical, historical researchers, the least), and in their focus on individuals versus institutions (laboratory studies tend by necessity to be individualistic, other forms of inquiry tend to be more flexible). Moreover, it should be emphasized that there is a great deal of intracategory variability. Laboratory studies of cognitive, affective, and small group processes differ greatly among themselves—in the types of tasks presented, the measures collected, the subjects used, and the hypotheses tested. Historical case studies vary dramatically in the thoroughness and comprehensiveness of the examination of evidence, the rigor of the hypothesis testing, and the systematic use of "focused comparisons" with other cases. Multivariate statistical studies of international conflict rely on an enormous range of data bases (attributes of nations, events data, capabilities data, political rhetoric), use a variety of analytic techniques to process these data, and test hypotheses derived from all the major levels of analysis considered earlier. In short, this chapter can but skim the surface of these vast research literatures.

Laboratory Experiments and Simulations

Researchers have used experiments and simulations to explore a variety of psychological and social processes that, with minimal imagination, can be seen as relevant to the problem of avoiding nuclear war. Voluminous experimental literatures exist on decision making in general (Abelson and Levi, 1985; Einhorn and Hogarth, 1981; Kahneman et al., 1982) and decision making under high-stress conditions in particular (Janis and Mann, 1977; Streufert and Streufert, 1978), attitude formation, persistence, and change (McGuire,

1985; Nisbett and Ross, 1980), and bargaining and negotiation processes (Pruitt and Rubin, 1986). The internal validity advantages of exploring basic processes under controlled laboratory conditions are well known. Investigators are in a position to make relatively strong inferences concerning determinants of the response class of interest. They can isolate the effects of particular independent (experimentally manipulated) variables, control for potential confounding variables, systematically assess interactions among independent variables, and test detailed models of the processes hypothesized to mediate relationships between independent and dependent variables.

Experimental research can obviously take many forms, from highly schematized payoff matrix studies (for example, mixed-motive games such as the prisoner's dilemma) to much more elaborate and involving scenarios (for example, the inter-nation simulations). Regardless of form, however, skeptics have been more impressed by the dissimilarities than the similarities between laboratory studies and high-level political decision making. The skeptics note the many differences between the almost always college student subjects and top policymakers: maturity, background, training, and concern for doing the best possible job (policymakers are presumably extremely motivated; college student subjects may approach their experimental assignments with a more casual attitude). The skeptics also note the existence of organizational and systemic constraints on actual policymakers that are not present in the laboratory. Policymakers, unlike experimental subjects, are embedded in complex networks of intra- and interorganizational accountability and competition. Finally, the skeptics contrast the brief duration of experiments (typically an hour or two) with the protracted course of most foreign policy deliberations (most crises last for days). In short, it is trivially easy to identify a plethora of threats to the generalizability of experimental studies, each threat a potential boundary condition on the applicability of elegantly demonstrated experimental phenomena to the complex, confusing, even chaotic world of foreign policy.

Although it is easy to point out ways in which laboratory experiments differ from the "real thing," it is much more difficult to specify when the differences make a difference. How does one begin to close the yawning generalizability gap between the artificial world of the laboratory and the "real" world of international conflict? Our contributions have not found a simple, all-purpose solution to this problem. In general, they proceed cautiously and rarely rely solely on experimental evidence to make an important argument. Consider the following three examples from *Behavior, Society, and Nuclear War:*

1. Holsti (this volume) refers to the large body of experimental work that has revealed a inverted-U relationship between stress and judgment and choice processes. He notes that very high levels of stress (threat to important values

and information overload) are generally disruptive of complex cognitive functioning. But Holsti also invokes other work—in organizational theory, history, and political science—to support claims concerning the effects of crisis-induced stress on the foreign policymaking process.

2. Stein (forthcoming) draws on laboratory research on bargaining processes to highlight the limitations of a pure-threat deterrence posture and to underscore the importance of drawing on alternative influence tactics (Chertkoff and Baird, 1971). Stein, however, draws even more heavily on historical case studies. Similarly, Stein notes laboratory research on motivated distortion in social judgment (Janis and Mann, 1977), but again puts even more weight on historical case studies that suggest leaders of challenger states sometimes engage in wishful thinking when assessing the chances of their successfully challenging the deterrence posture of a status quo power.

3. Druckman and Hopmann (this volume) refer to a variety of experimental studies for potential insights into international negotiation—including work on the effects of reciprocity strategies (Wilson, 1971), the effects of role reversal (Walcott et al., 1977), and negotiator responsiveness (Pruitt and Rubin, 1986). But, they are willing to draw inferences for the conduct of foreign policy from these literatures only in the presence of converging evidence from historical studies of international conflict management.

Mass Surveys of Public Opinion

Surveys do not permit the kind of confident causal inferences possible in experiments. The representative sample survey is, however, the method of choice for investigating a large class of questions relevant to nuclear war. Surveys allow us to assess the breadth and depth of popular support for particular foreign and defense policies, attitudes toward key political events and personalities, and the degree of lability-stability in public sentiment on these questions. Insofar as public opinion constrains, and perhaps occasionally even drives foreign and defense policy, surveys can shed light on important inputs into the policymaking process.

Like experimentation, survey research can take diverse forms. Researchers can focus on interrelationships among variables at one time or across time; they can focus on a narrow or broad range of political issues; they can focus on the general public or political elites or both; they can assess the impact of major events on public opinion; they can focus on the complex feedback relationships that appear to exist between public opinion and public policy; they can even embed question-wording experiments into surveys to assess the susceptibility of public opinion to linguistic manipulation. These methodological variations illustrate the flexibility of survey research and the range of issues that can be addressed.

A variety of by no means intuitively obvious findings have emerged from

survey research on U.S. national security attitudes (see Russett, this volume). Public opinion on such issues does not appear to obey the same rules of ideological constraint as elite opinion (Converse, 1964). Many Americans in the 1980s, for example, do not feel it is inconsistent to support both a "freeze" on nuclear weapons and President Reagan's Strategic Defense Initiative (SDI). Citizen support for particular policies also often depends greatly on the linguistic packaging of the policies. The insertion and deletion of affect-laden terms (the presidency, communism, national defense, the Soviet Union, etc.,) can have large effects on response distributions (Schuman and Presser, 1981). Similarly, so can the relative emphasis on the "gain" and "loss" components of a policy proposal. Questions that make one's own losses salient (for example, the missiles we are required to dismantle) or the other side's gains salient (e.g., the weapons systems they get to keep) will obviously evoke less popular support than questions that make the other side of the trade-off equation salient (Kahneman and Tversky, 1979).

It is especially important from a policy standpoint not to underestimate the magnitude and pervasiveness of question-wording effects. Pundits commonly speak of public opinion on issues such as the SDI or the nuclear freeze as though it were a straightforward matter to obtain a single-point estimate that, within confidence intervals set by sampling variability, reflects how the American people think about key national security issues. Matters are not, however, so simple. One can sometimes obtain very different estimates of public sentiment by posing only subtly different versions of the same questions. Not surprisingly, political partisans often disagree sharply over what constitutes the fairest phrasing of questions, with judgments of fairness suspiciously highly correlated with degree of public support elicited for preferred positions.

Unfortunately, there is no theory, accepted procedure, or standard approach for adjudicating disputes over question bias and fairness. To be sure, numerous methodological tactics might be deployed here. One might try to develop politically neutral versions of questions; one might ask many (biased) versions of the "same" question; one might assess the stability of opinion in response to "neutral" questions by assessing how much people are willing to change their attitudes when confronted with particular challenges or counterarguments. In each case, however, there is always room for the political prejudices of the investigator to contaminate the results. One can never be sure that one did not phrase the argument on one side more persuasively than the argument on the other. Our knowledge of how to construct unbiased samples from the general population of potential respondents has advanced much more rapidly than our knowledge of how to construct unbiased samples from the conceptual population of potential questions.

Quantitative Studies of Historical Data

Many investigators have attempted to extend the empirical grasp of behavioral and social science by quantifying archival data and then using these data to test hypotheses concerning correlates or determinants of war. This methodological approach is not linked to any specific level of analysis. Efforts at quantitative hypothesis testing using historical data occur at a psychological level of analysis (for example, content analysis studies of policy deliberations and diplomatic communications: Axelrod, 1976; Holsti et al., 1969; Tetlock, 1985a), at the level of the nation-state (for example, Singer's [1980] correlates of war project) and at the level of the international system (for example, Bueno de Mesquita's [1985] efforts to test his theory concerning the necessary conditions for the rational initiation of war by developing quantitative systemic indicators of the balance of power).

The advantages of quantification are well known: investigators must use explicit, consistent, and public rules for attaching numerical scale values to observations. And the disadvantages of quantification are equally well known, in large part because of these very strengths. In their efforts to represent complex theoretical constructs with simple empirical indicators, quantitative researchers often make assumptions about their data that many skeptics regard as politically naive at best and ridiculous at worst.

It is no simple matter, for example, to gauge the importance of "perceptions of capability" in the policy deliberations of decision makers prior to World War I. One possible indicator is the frequency with which policymakers discuss such issues in the archival records (Holsti et al., 1969), but the absence of such discussion is far from compelling evidence that perceptions of capability played no role in the crisis decision-making process (Jervis, 1970). Decision makers may not have bothered to express such concerns because they felt the relevant information was too widely known or obvious, or the archival record may just be incomplete. It is also no simple matter to gauge the impact of internal conflict within a state on its propensity to go to war. There are many possible indicators of internal conflict (civil disturbances, guerrilla warfare, attempted coups, labor unrest, inflation, unemployment, intensity of ethnic, religious, or political factionalism), indicators that are frequently only weakly intercorrelated. It is even very difficult to achieve consensus on superficially straightforward judgment calls such as who initiated or won a war. Critics of Bueno de Mesquita's (1981) expected utility theory argue that his tests of the theory rest on a number of dubious historical classifications (Majeski and Sylvan, 1984). For instance, Bueno de Mesquita operationally defines the initiator of a war as the first state to engage in sustained combat on the opponent's territory—a definition that adequately

covers many cases but seriously oversimplifies others (for example, it is at least debatable that Israel was the sole initiator of the 1967 Six Day War). And it strikes many observers as odd to classify Serbia as the sole victor in World War I or Poland as a victor in World War II (see Levy, this volume).

The difficulty of interpreting many quantitative historical indicators does not, of course, mean that it is fruitless to try. It does, however, help to explain why it has proven so hard to identify a set of predictively powerful and uncontroversial laws of international conflict. The explanatory variables are often extremely complex, context bound, and resistant to precise, standardized operational definitions. Partly because of these concerns, many researchers interested in international conflict eschew quantitative methods in favor of more flexible, case-specific, qualitative methods.

Case Study Methods

Although case study methods are perhaps the least scientifically prestigious of the approaches to generating knowledge considered here, such methods play a critical role in research on the origins of international conflict. As George (1979) and Eckstein (1975) have argued, case studies can both build on and enrich experimental and statistical approaches to the study of political processes. Case studies can be used, for example, to stimulate new lines of theorizing (for example, George and Smoke's [1974] comparative historical studies of deterrence suggested a variety of hypotheses concerning the conditions under which threats of force are likely to elicit desired reactions from other states), to provide detailed, qualitative evidence that phenomena documented by experimental or statistical research do indeed occur in specific historical situations (for example, Jervis' [1976] widely acclaimed use of diplomatic history to show that policymakers fall prey to cognitive biases and errors documented by experimental psychologists), and to cast doubt on the general validity of particular generalizations or lawlike claims (for example, Lebow's [1981] use of case studies to challenge the deterrence theory claim that the most important cause of international aggression is the perception of the leadership of "challenger" states that "status quo" states lack the resolve or capability to defend commitments). In short, case studies can advance knowledge in diverse ways.

The contributors to this series frequently draw on case study evidence to formulate, support, and qualify theoretical claims. This reliance on case study evidence reflects, in large part, a recognition that experimental and statistical studies—notwithstanding their many strengths—fail to capture much of the subtlety and complexity of the actual conduct of foreign policy.

Case studies complement quantitative, variable-centered research by pro-

viding qualitatively rich and contextually detailed descriptions of the lives and events that we seek to understand. The explanatory goal is no longer the creation of statistical models that account for as much of the variance across cases as possible; the goal is the creation of conceptual models that organize the disparate themes and strands of meaning that run through particular historical events. Thus, if one wants to understand the dynamic ebb and flow of U.S.-Soviet talks on intermediate-range nuclear forces in Europe, it may help to be aware of experimental and quantitative field research on bargaining and negotiation. There is no substitute, however, for "thick description" of the specific events of interest (Geertz, 1973; Talbott, 1985).

Thick description entails much more than a complete behavioristic account of an event (who said what to whom, when, where, and how?); it requires an interpretive account of the "multiplicity of complex conceptual structures, many of them superimposed upon or knotted into one another" (Geertz, 1973:10) that give meaning and structure to foreign policy. One needs, for example, to appreciate the intricate sociopolitical systems within which arms control negotiations are embedded. One needs to be aware of the crosscutting personal and political ambitions and rivalries within the negotiating teams, the degree of autonomy granted the negotiating teams on specific issues by their respective governments, how the limits of that autonomy were negotiated and are now understood, the domestic political constraints within which key governmental decision makers must operate, intra- and interalliance politics, the views of key decision makers on the types of concessions it is realistic to hold out for, and so forth (Talbott, 1985). Any given act within a negotiation session is open to interpretation at any one or combination of these various levels of analysis—a state of affairs that strains the open-mindedness even of investigators who are exceptionally tolerant of ambiguity.

To achieve highly nuanced case descriptions of this sort, investigators must often proceed more by feel and improvisation than by plan and research design. They must sift through often complex and contradictory archival records and through interview protocols with policymakers that yield difficult to disentangle mixtures of candid revelation, distorted recall, and self-serving rationalization. As a result, it is extraordinarily difficult to be explicit and systematic about standards of data collection and interpretation in case studies. It is also extraordinarily unlikely that even two investigators working from exactly the same data set will reach identical conclusions. Sometimes the disagreements will revolve around differences in emphasis. The investigators will attach different weights to different data sources. Sometimes the disagreements will be more fundamental. The investigators will reach opposite conclusions about the necessity or usefulness of drawing on a particular level of analysis to make sense of the events in the case history.

This apparent lack of methodological rigor is at the heart of many objections to case study approaches. Verba (1967) notes, for example, that case studies do not add up easily. Although each study may be beautifully written and elegantly organize a wide range of historical facts, it is rarely possible—because of idiosyncrasies in methods of gathering and interpreting data—to derive reliable and valid theoretical statements. Without well-defined and consistent standards of evidence and procedure across cases, there is no clear way of determining whether the variables are being measured on the same "scales" or, for that matter, whether the same variables are even being measured. There are also relatively few checks on the intrusion of error and bias into the research process. It is hard to say how much emphasis researchers have put or should have put on particular items of information in drawing particular conclusions.

These differences in emphasis can, moreover, be theoretically consequential. For instance, whether one views an ambiguous historical case as consistent or inconsistent with "deterrence theory" may ultimately hinge on the credibility one believes the leadership of the challenger state attached to warnings or threats from the status quo state. This judgment, in turn, hinges on a rather precise reconstruction of the expected utility decision calculus of the leadership of the challenger state from rather crude historical clues (see, for example, the exchange between Lebow [1987] and Orme [1987]). How seriously did the challenger view certain statements by certain government officials of the status quo power? Did the challenger under- or overestimate the military capabilities or alliance cohesiveness of the status quo power? How seriously did the challenger view the military preparations of the status quo power? Did the challenger under- or overestimate the significance of domestic political opposition to deterrence within the status quo power? Analysts of the historical case must try to piece together answers to such questions from often fragmentary and inconsistent archival records. Confronted with such a complex and unstructured task, it would be surprising if even methodologically self-conscious investigators did not occasionally fall prey to the cognitive tendency to give more weight to hypothesis-consistent than hypothesis-inconsistent evidence in drawing conclusions from historical records (Nisbett and Ross, 1980).

Defenders of case study approaches have a number of possible defenses to these objections. George (1979) and Janis (1982) forcefully argue that many of the methodological and inferential flaws commonly linked to case study approaches are by no means intrinsic to this genre of research. Janis' (1982) comparative case studies of groupthink in foreign policy deliberations, and George and Smoke's (1974) comparative studies of deterrence in international relations demonstrate that investigators can sometimes capitalize on the dis-

tinctive strengths of both quantitative, nomothetic studies (for example, explicit statements of theoretical hypotheses and explicit efforts to test those hypotheses against the evidence) and qualitative idiographic studies (for example, sensitivity to the unique historical circumstances of each case). From this perspective, case studies are a major methodological means of advancing the search for general laws or patterns underlying international conflict. Case studies serve the same ultimate epistemic goal as laboratory and statistical field studies (subsuming new observations under Hempel-like covering laws or identifying exceptions to these covering laws that require developing more complex or contingent theoretical generalizations).

Other defenders of case study approaches might mount even more radical challenges to quantitative, variable-centered research. Well-executed case studies serve as sobering reminders of just how difficult it is (perhaps impossible) to know whether a given covering law applies in a given setting. To be sure, it is possible for quantitative content and event analysts to develop systematic schemes for coding negotiation behavior (Pruitt and Lewis, 1975; Stephenson et al., 1977), political rhetoric (Holsti et al., 1969; Tetlock, 1985a), and foreign policy actions (Leng and Wheeler, 1979; Leng, 1983). But these efforts at quantification are—from a radically idiographic perspective—profoundly misguided. Efforts to develop coding categories that place superficially very different acts in the same theoretical categories encourage investigators to downplay, even ignore, highly context-specific components of meaning. Harsh words exchanged at the negotiation table may reflect a personal animosity, a secret joke at the expense of other participants or higher-ups, a carefully orchestrated exchange designed to manipulate the press of other countries, a calculated effort by one or both parties to sabotage the talks, and so on. The possibilities are virtually endless and only diligent attention to the particulars of each case can clarify which interpretations make sense of the facts—indeed, can clarify which facts need to be made sense of. Mechanically coding what is said on the basis of strictly syntactic or semantic criteria (hostile-friendly, responsive-unresponsive, simple-complex, etc.) makes it possible for quantitative comparative investigators to achieve intercoder reliability, but only at the expense of validity—of doing extreme violence to the elaborate networks of context-specific understandings that participants in political settings have worked out among themselves.[5]

Formal Mathematical Models

A final method of generating knowledge about the sources of international conflict deserves mention. Whereas practitioners of case study methods emphasize the importance of "thick description" of the historical contexts within

which events are embedded, practitioners of formal modeling emphasize the importance of precise understanding of the underlying structural processes that drive historical events—underlying processes that can be most parsimoniously and effectively described by game theory (Shubik, forthcoming), differential equations models (Intriligator and Brito, 1984), or other advanced mathematical tools.

Although work in this analytic tradition can take many forms, game theory analyses clearly predominate. As Shubik (forthcoming) notes, game theory provides a "formal tool" for exploring what happens when rational, goal-oriented individuals interact with each other in particular environments. These environments can be defined by payoff matrices that specify the outcomes each party can expect given the response options that both have chosen. Game theorists argue that it is possible to reduce the enormous complexity of international conflicts to a finite set of mathematically well-defined games such as the "prisoner's dilemma" and "chicken" (see Shubik, [forthcoming] for more detail). War, in this view, arises not because of the cognitive shortcomings of leaders or the political shortcomings of nations but because of the incentives and disincentives that are built into payoff matrices that, in turn, capture the essence of the international predicaments confronting national leaders.

By way of illustration, one game theorist, Brams (1985:145), has argued that many of the most intractable issues that divide the United States and Soviet Union can be understood as products of the "unforgiving nature" of certain two-person nonzero-sum games. He attempts to model nuclear deterrence, for example, with the game of chicken (see also Schelling, 1966). The basic task of a player who desires to deter an adversary is to make the choice of aggression sufficiently unattractive—through threats of retaliation—that the adversary will refrain from undertaking the act. The key difficulty with this strategy, especially in a world of mutually assured destruction, is the shared knowledge that the deterrer will inflict grievous harm on himself or herself as well as his or her adversary if he or she actually executes the threat of retaliation. An important task for game theorists then becomes the mathematical solution of the difficult problem of identifying an "optimal compromise" (Brams, 1985:147) between the need for deterrent threats that are both effective and credible. Brams (1985) also tries to model the nuclear arms race by employing the prisoner's dilemma game. Both sides in this game would be better off cooperating, he notes, but fear of exploitation keeps them in competition. An important task for game theorists here becomes the identification of a strategy of conditional cooperation in which each side has the monitoring capability to ensure that the other side cooperated when it said it would.

Game-theoretic analyses sometimes yield startlingly simple but logically compelling conclusions. For instance, Axelrod (1984) conducted a computer

simulation study that pitted a large number of expert-recommended strategies for coping with the prisoner's dilemma against each other. The simplest submission—tit-for-tat—won the most points—a submission that, Axelrod argues, embodied four critical strategic attributes (it was nice, clear, forgiving, and retaliatory).[6] Jervis (1978) subjected the widely used concept of security dilemma to detailed logical analysis to probe the conditions under which competition or cooperation is most likely to occur in prisoner's dilemma types of international situations. Policies that decrease the potential losses of unrequited cooperation or the potential gains of unilateral defection appear most likely to encourage stable mutual commitments to the otherwise unstable "cooperate-cooperate" cell of the payoff matrix. Jervis (1988) also makes an observation of special interest, at a time when serious consideration is being given to a new technological generation of antiballistic missile systems. He notes that cooperation in prisoner's dilemma types of international environments is more likely, to the degree that defensive military systems can be readily distinguished from offensive ones.

Investigators have applied other formal modeling approaches as well to the study of international conflict. Most important perhaps have been the efforts to develop differential equation models of arms races and to identify the conditions under which arms races do and do not lead to war. Early work that suggested arms races are inherently destabilizing appears to have been superceded (Richardson, 1960). Intriligator and Brito (1984), for instance, have argued on the basis of their interesting mathematical model of competitive military buildups that arms races can lead to war *or* peace, depending on the initial configuration and balance of forces and on the nature of the race (whether "qualitative" advances in weapons technology such as equipping missiles with multiple independently targetable re-entry vehicles (MIRVs) occur). The critical determinant—of whether arms buildups (or, for that matter, arms reductions) increase the likelihood of war—is whether the two sides have moved into a "force space" in which one side can successfully attack the other. It should be noted, however, that such analyses as these are compelling only insofar as one is willing to grant the empirical reality of the underlying mathematical assumption. There are usually solid grounds for skepticism. For example, from the point of view of Intriligator and Brito's (1984) model, it makes no difference whether one passes through a region of instability (breakdown of mutual deterrence) as a result of an arms buildup or as a result of compliance with an arms reduction agreement; from a psychological and political point of view, it may make a great deal of difference.

Formal modeling, especially game-theoretic, approaches have attracted criticism from a variety of quarters. The most frequent criticism is paradoxically directed at what many defenders view as the greatest strength of formal

models: the deductive simplicity and elegance of the formulations. To critics, this elegance bespeaks lack of psychological and political realism. People often lack clear goals, misperceive each other's actions and intentions, and miscalculate what is in their own best interest. Nations often send off unintended signals and respond as much to internal political necessities as to external systems of incentives. In short, skeptics can challenge the generalizability and relevance of game-theoretic "solutions" to international conflict in much the same way that they can challenge the generalizability and relevance of experimental research findings—by pointing to potentially powerful variables at work in the international environment that were not taken into account in the original research.

The methodological challenge is to make the connection between the austere formalisms of game theory and the messy world of international conflict. Oye (1986) offers a fascinating example of how this might be done. A series of six historical case studies was commissioned to test the impact of three game-theoretic structural variables on international conflict: the mutuality of interest (cooperation increases as a function of the relative strength of the payoffs to cooperate versus compete), the shadow of time (cooperation increases as the temptation to obtain short-term gains from competition decreases), and the number of players (the fewer players, the more cooperation). The results reveal both the strengths and limitations of a purely game theoretic approach. In Axelrod and Keohane's (1985:227) words, the three causal variables deduced from game theoretic analyses "help us to understand the success and failure of attempts at cooperation in both military-security and political-economic relations." They add, however, that the structural variables, either separately or jointly, are not sufficient for cooperation. There is a multitude of impediments to cooperation that are, at least at present, extremely difficult to capture in formal game theoretic models—including ideological and cognitive variables, organizational and bureaucratic variables, and domestic political variables (variables embedded within levels of analysis reviewed earlier in this chapter).

The Quest for Linkages

Readers of this series should expect both theoretical and methodological diversity. There is no single, unified theory of international conflict; there is, instead, a continuum of theoretical perspectives, ranging from the micro (psychological) to the macro (systemic), within which investigators can formulate hypotheses and conduct research. And there is no single set of methodological guidelines for studying international conflict; there is, instead, a

broad range of methods that, depending on the level of analysis and the type of problem under investigation, researchers are likely to find more or less useful.

How should we react to this confusing plurality of theoretical and methodological perspectives? One possibility is that the confusion is temporary—a reflection of the immature ("preparadigmatic") state of theoretical and methodological development in research on international conflict. It is only a matter of time before one level of analysis and set of research methods come to dominate inquiry. One version of this "waiting for a paradigm" thesis is microreductionist. Investigators will eventually be able to show that macrophenomena such as wars among nations can be best understood in terms of basic (probably experimentally demonstrated) laws of individual behavior. The epistemological mirror image of this approach is, of course, macroreductionist (what Greenstein [1975] has aptly called the "actor dispensability thesis"). Investigators will eventually be able to explain conflict among nations in terms of the operation of institutional, cultural, domestic political, economic, or systemic forces, with recourse to only minimal assumptions concerning the nature of individual decision makers.

Another possibility is that this plurality of theoretical and methodological perspectives will be with us for a very long time indeed—a reflection not so much of the immaturity of the disciplines as of the complexity of the subject matter. From this perspective (and it is this perspective that has guided chapter selection for this series), it is unwise to assume that the causal nature of the micro-macro relationship can be known in advance or that this relationship is always and everywhere the same. A more reasonable starting point is to assume that micro- and macroprocesses typically interact to shape decisions bearing on war and peace, with the degree of linkage and exact balance of causal forces shifting from time to time and under different conditions.

This open-ended interactionist position suggests that searches for a fundamental or unifying level of analysis are misguided. Rather than seeking to fit all research efforts into a common reductionist mold, we should be content with: (1) looking for linkages across levels of analysis (ways in which different levels complement and mutually enrich each other); and (2) looking for linkages across research methods (assessing the degree to which practitioners of very different methods of research reach compatible or contradictory conclusions).

Theoretical Linkages

Advocates of different levels of analysis often seem to speak in different data languages—languages that are so different that they cannot be translated

into each other without egregious loss of meaning. Balance-of-power theorists, for instance, are not interested in using or testing detailed process models of individual decision making; they feel that it is possible to achieve self-contained, internally consistent, and predictively powerful accounts of international conflict by relying on purely systemic indicators. Investigators in this tradition design studies to test which configurations of systemic variables best explain when, where, and why war breaks out (variables from other levels of analysis simply drop out of the empirical picture).

Although it is true that each level of analysis reviewed earlier has generated its own distinctive, self-contained research literature, it is also true that the distinctions between (and among) levels of analysis are not nearly as neat and tidy as academic writers sometimes imply. Levels of analysis "interpenetrate." One can make a strong case that variables operating at a micro- (psychological) level of analysis rarely directly determine policy outcomes; micro-level processes are constrained, shaped, and perhaps sometimes even transformed by the social systems within which individual decision makers must work and by the structure of the problems that they must confront. Conversely, one can make a strong case that our understanding of when and how macroprocesses shape policy outcomes would be much enhanced by drawing on more realistic assumptions concerning the nature of the individual decision maker.

Let us consider a few examples of how micro- and macroapproaches to the study of war and peace might be brought together. As noted earlier, it is frequently argued that how individual decision makers respond to policy problems reflects the operation of internal psychological processes—for instance, the tendency to rely on simple judgment and choice heuristics to reduce cognitive strain and the tendency in evaluating options to encode possible outcomes as gains or losses from a neutral reference point (Kahneman et al., 1982). Exactly how these response predispositions are expressed in a specific foreign policy setting may depend enormously on the surrounding social-political context. Many of these macro constraints are so obvious that they hardly need to be specified. The "content" of thought—the policy options considered (for example, arms control proposals), the consequences contemplated to arise from each option (for example, the impact on different parties' perceptions of the balance of power), and the bureaucratic and political constituencies to be placated—is largely dominated by the policymaker's perception of macrolevel variables.

It would be misleading, however, to conclude that the disciplinary division of labor is quite so simple, with cognitive psychologists specifying the abstract information processing rules used for interpreting events and making choices, and political scientists and historians specifying the "belief-content"

on which those abstract rules operate. It is quite conceivable—available research indicates even likely—that the processing rules themselves can change as a function of the social-political environment. Much depends—as work on group dynamics and bureaucratic politics suggests—on the role and accountability relationships that exist among key decision makers.

These role and accountability relationships link individual decision makers to social systems and can increase or decrease, for example, the complexity of the cognitive strategies that decision makers use to make sense of the world (March, 1978; Halperin, 1974; Janis, 1982; Tetlock, 1985a; Weick, 1979). Much also depends on the nature of the problem confronting the decision makers and exactly how the problem is presented to them (Einhorn and Hogarth, 1981). For instance, although decision makers generally find trade-off reasoning aversive and tend to define situations in ways that deny or minimize trade-offs, there seem to be certain institutional and strategic environments in which trade-offs are so starkly obvious that they have become, in effect, undeniable. Steinbruner (1987:535) has advanced such an argument with respect to the command and control of nuclear weapons. The extreme destructiveness and rapid timing of these weapons has forced the superpowers to confront "unavoidable conflicts among fundamental objectives"—objectives such as maintaining control over one's nuclear forces in a crisis and retaining the ability to respond to a massive first strike. The "solutions" to profoundly difficult trade-offs of this sort are reflected in the institutional procedures that the superpowers have evolved to plan and direct their strategic operations. It would be extremely difficult to argue for a change in these institutional procedures without simultaneously acknowledging the importance of the major conflicting values in the trade-off equation.

Knowledge of response dispositions that exist at the microlevel is then helpful, but rarely sufficient. One can, moreover, make almost an identical argument with respect to knowledge of macrolevel processes. Macroprocesses surely constrain, sometimes sharply, the range of conceivable outcomes of policy deliberations. National leaders sometimes feel that "their hands are tied," that they have no choice (save resignation or waiting to be ousted from office) but to act in certain ways. President Kennedy reportedly felt that he had to succeed in removing Soviet missiles from Cuba or face impeachment (Allison, 1971). Macrotheorists have yet, however, to provide persuasive empirical demonstrations that they can reliably predict specific policy outcomes from the values of macrolevel variables. Too many exceptions exist. Policymakers sometimes resist pressures to represent narrowly defined bureaucratic interests when they feel a larger national objective is at stake; political leaders sometimes court disaster by advocating policies that antagonize important constituencies; national leaders do not always decide to

go to war when the "expected utility of war" (as gauged by the sorts of systemic indicators used by Bueno de Mesquita, [1985]) is positive, even highly positive.

Macrocauses of war do not seem to operate the same way on different national leaders and in different situations. As soon as we come into contact with the historical record, simple bivariate hypotheses derived from macrotheories need to be qualified (Greenstein, 1975; Levy, this volume). The complexity arises, in part, because we live in a multivariate macroworld—a world in which many causes at the macrolevel are interactively shaping policy. A policymaker may refuse to yield to pressures from one constituency because he or she is under even greater pressure from another. Or the leader of a hegemonic state may refrain from preemptive war against a rapidly rising challenger state because of the looseness of his or her own state's alliance structure or the tightness of the alliance structure of the challenger.

The complexity also arises because we understand quite poorly exactly how: (1) macrovariables constrain processes at work at a microlevel; (2) microprocesses aggregate to produce macro-outcomes. For instance, systemic theories (the most macro of the macrotheories) assume that, in order to survive in an anarchic international system, states must attach "primacy to their security interests" (Levy, this volume). This assumption tells us very little about how risk-seeking or risk-averse states will be in their pursuit of power in different situations. Macrotheories could benefit in this regard by drawing on some empirically well-validated propositions concerning individual decision processes. For example, prospect theory (Kahneman and Tversky, 1979) leads one to expect leaders to be much more willing to take large risks in order to avoid major losses in national power than in order to expand national power. This tendency to loss aversion may help to explain why threats to the balance of power have been so strongly associated with the outbreak of general wars in the last five centuries. Wars frequently arise in such situations as a result of dominant states fearing loss of control and launching preemptive wars or as a result of weaker states coalescing to prevent a would-be hegemonic power from imposing its will on them.[7]

The tendency to loss aversion may also help to explain some intriguing patterns in public opinion data on support for national security policies. It has been argued, for instance, that the public is most willing to back hardline policies when "national pride" has been wounded (McClosky, 1967)—an explanation that has been invoked, albeit in post hoc fashion, to account for the receptiveness of the German public to nationalistic appeals in the wake of the Versailles Treaty and for the surge in American public support for defense spending in the aftermath of the Vietnam War, OPEC oil embargoes, and the Iranian hostage crisis. It has also been noted (Russett, this volume) that

American public support for defense policies tends to increase when those policies (SDI) are presented as means of avoiding potentially catastrophic losses (the destruction of U.S. cities). Avoiding an easily imaginable disaster seems to be a much more psychologically compelling objective than reaching a difficult-to-understand and perhaps even more difficult-to-justify arms control agreement.[8]

Explicating the microprocesses that link macrophenomena to foreign-policy decision making is a profitable but surprisingly underutilized way of exploring linkages among levels of analysis. From a psychological point of view, systemic theories are often woefully underspecified (Simon, 1985). It is not enough simply to posit the existence of rational national actors who maximize their interests within the constraints of the international system. One can derive very different predictions about the effects of important systemic variables depending on the auxiliary assumptions that one makes about the subjective probabilities or beliefs and utilities or preferences of the policymakers involved. Three examples—all highly relevant to the current geopolitical scene—must suffice to make this critical theoretical point.

>   1. *Is nuclear proliferation destabilizing?* Enormous concern has been expressed over the slowly but inexorably expanding number of nuclear powers. One can construct plausible systemic arguments that this concern is either justified or unjustified. One can argue that nuclear proliferation reduces the likelihood of war by inducing caution, increases the likelihood of war by increasing the incentives for preemption, or has no effect one way or another (because the former two effects cancel each other out). Which outcome one predicts hinges on the assumptions one chooses to make about the belief systems and utility functions of key national decision makers. (In extreme cases, it may even be necessary to redefine what is customarily meant by "rational." What happens, for example, if nuclear weapons fall into the control of messianic religious or political leaders who attach much higher value to destroying their enemies than they do to insuring their own survival?)
>
>   2. *Is parity destabilizing?* Organski and Kugler (1980) have challenged the widely held view that approximately equal distributions of power among major states are conducive to peace. They have argued that parity in power is actually destabilizing because it tempts each side to believe that it has a reasonable chance of winning. They have argued, moreover, that parity is particularly dangerous when the balance of power is in flux. Which position one takes depends in large part on one's assumptions about the accuracy with which policy elites can appraise shifting military–technological–economic balances of power. The Organski and Kugler position seems to leave more room for cognitively or motivationally driven forms of misperception than the traditional realpolitik position (Stein, forthcoming). Which position one takes may also depend on one's assumptions about key decision makers' attitudes toward risk

and uncertainty. The Organski and Kugler position seems to imply a greater willingness to take risks than the traditional realpolitik analysis.

3. *Are alliances destabilizing?* Levy (this volume) notes that there has been much controversy concerning when alliances increase versus decrease the likelihood of war. There is no single answer to this question. The impact of alliances depends almost certainly on the degree to which alliances simplify the calculations of would-be aggressors (uncertainty reduction), make war more or less attractive (by affecting the subjective probability of success), and make war seem more or less likely (by affecting perceptions of the hostility of the other side). Systemic theories can make predictions on this key geopolitical issue only by assigning implicit or explicit causal weights to each of these components of the individual and collective decision-making process.

Methodological Linkages

Just as communication across levels of analyses can be difficult, so too can communication across research methods. The difficulties are not, however, insurmountable. I have already mentioned several examples of multimethod convergence. Sometimes very different methods of inquiry yield compatible conclusions. Thus, laboratory studies of judgment and choice, quantitative content analyses of archival records, and qualitative case studies all point to a widespread tendency for decision makers to rely on simple, low-effort heuristics in interpreting new events and choosing among courses of action (Axelrod, 1976; George, 1980; Jervis, 1976; Nisbett and Ross, 1980). The same three categories of method—plus some game theoretic and computer-simulation work (Axelrod, 1984)—also point to a common conclusion concerning the relative effectiveness of different influence tactics in mixed-motive games. In general, some form of tit-for-tat (reciprocity) strategy is more effective than either bullying or appeasement as a method for achieving mutually beneficial compromise agreements (George et al., 1971; Leng and Wheeler, 1979; Pruitt and Rubin, 1986; Snyder and Diesing, 1977). Many additional examples of convergence could, moreover, be cited (Holsti, this volume; Stein, forthcoming).

The notion of seeking out multimethod convergence is deeply entrenched in the behavioral and social sciences (see Campbell and Fiske, [1959] on multiple operationism), so it should not be surprising to see the idea surface in a research domain where there is so much uncertainty concerning both what needs to be measured and how the measurement process should proceed. It seems only prudent not to put all of one's theoretical "eggs" in one methodological "basket"—to recognize that different methods often have complementary strengths and weaknesses and to look for patterns of convergence in the findings that emerge from applications of these methods. And it is reassur-

ing that the search for multimethod convergence has occasionally been successful. Investigators working with quite different theoretical concepts and very different methodological tools have sometimes arrived at surprisingly similar conclusions.

There is, then, some cause for optimism that many of the emergent generalizations discussed in *Behavior, Society, and Nuclear War* are not method specific. Implementing a multimethod research strategy is not, however, a simple task, for a number of reasons. Part of the problem is the difficulty of determining whether multimethod convergence has indeed occurred; the other part of the problem is the difficulty of deciding what to do when one concludes that multimethod convergence has not occurred—when different methods yield inconsistent, even contradictory, results.

There is no fixed, objective rule for solving these problems. Consider, for instance, the difficulties that arise in assessing whether experimental research on mixed-motive games really converges on the same conclusions as qualitative and quantitative studies of the historical record. A "fair-but-firm" reciprocity strategy may have a precise operational definition in the laboratory, a fuzzier, more open-ended operational definition in quantitative event analysis studies, and a highly context-specific operational definition in historical case studies (a definition anchored in politically controversial assumptions about the perceptions and goals of specific actors at specific times in the flow of events). Similarly, judgmental biases such as belief perseverance typically have precise meanings in experimental studies but are notoriously resistant to precise or consensual definition in actual foreign-policy settings. (To what extent, for example, should observers of the Soviet Union have changed their minds after learning of the invasion of Afghanistan or after learning of the Soviet withdrawal?) The qualitative diversity of research methods makes it extremely difficult to determine whether practitioners of different methods are truly studying the same underlying phenomenon or whether we (the reviewers of interdisciplinary literatures) are imposing a false unity on these diverse research efforts. It is possible, as noted earlier, to have spurious multimethod convergence—to fail to recognize that superficially similar phenomena actually arise as a result of the operation of fundamentally different causes (for example, overconfidence in the validity of a simple historical analogy may arise not as a result of reliance on simple cognitive heuristics documented in laboratory work but rather from political pressures to appear "firm" to particular constituencies).

If identifying multimethod convergence poses problems, so too does interpreting multimethod divergence. When two different methods yield different conclusions, one confronts a plethora of interpretive options. One might question the usefulness of one method for testing a particular hypothesis or

class of hypotheses. Thus, it could be argued that multivariate correlational studies are, for various reasons, just not as useful for exploring linkages between domestic conditions and foreign policy as comparative case studies (see Levy, this volume). Or, it could be argued that qualitative forms of content analysis are more useful than quantitative techniques for identifying subtle shifts in the thinking of key national leaders or for predicting shifts in government policies (George, 1959).

Another interpretive option is to concede that both methods are useful for testing a given hypothesis and to argue that the different results have arisen as a result of the operation of moderator variables to which one method is more sensitive than the other. Thus, comparative case studies of decision making may be better equipped than laboratory studies to identify institutional and political boundary conditions on the expression of cognitively rooted judgmental biases (thus helping to explain why decision makers do not exhibit such biases in certain cases). Neither method, from this standpoint, is yielding trivial or artifactual results. The two methods, in conjunction, help to reveal the range of circumstances under which the hypothesized information processing biases hold up.

The key point is that how one decides to weight data from different methodological sources is ultimately a judgment call. There is no integrative set of guidelines that tells us when to pay special attention to, and when to ignore, results from particular research methods. And it is misleading to think of such decisions as purely methodological. Such decisions ultimately rest on implicit or explicit theories concerning the causal mechanisms that produce both regularities and irregularities in the data. Methodological and theoretical choices are, as we have seen before, tightly linked.

Concluding Remarks: The Quest for Policy Relevance

There are obviously many gaps and inconsistencies in the research literatures from which we have drawn in these volumes. But assume, for the sake of argument, that behavioral and social science research on international conflict had advanced much further than it now has. Assume that we possessed an integrative theory of international conflict that specified—with reasonable precision—how processes from different levels of analysis interact to shape policy outcomes. Assume that we also possessed broad interdisciplinary consensus on the usefulness of different research methods for testing different aspects of this integrative formulation. Would we then be in a position to offer authoritative advice on how to avoid nuclear war?

The answer is still not an unqualified "yes." Although such an integrative

theory would be enormously useful for organizing our thinking about problems of managing international conflict in general, it would not satisfy the needs of policymakers for guidance in coping with the myriad of specific real-world problems created by the introduction of nuclear weapons into international politics in 1945 and by the complex evolution and proliferation of such weapons and their delivery systems in the intervening 43 years. At their best, behavioral and social science theories yield conditional generalizations of the form: "Under circumstances x, y, and z, this type of intervention is likely to have these effects and under this other set of circumstances, the same intervention is likely to have this other set of effects." Such advice falls considerably short of telling policymakers whether it is a good idea to proceed with the development of a new weapon system or to accept a particular arms control proposal. Such advice falls short largely because it begs the question of how one determines whether the preconditions for adopting a given strategy have actually been met in a given situation. For example, it is one thing to claim that a firm-but-fair reciprocity strategy usually works better than alternative strategies (pure threat or appeasement) in promoting mutually advantageous solutions to conflicts of interest; it is quite another thing to claim that a reciprocity strategy is most appropriate in a particular political context. In Verba's (1967:116) words, "Generalizations fade when we look at particular cases." It is necessary to take into account the many circumstances unique to the case at hand, each circumstance not fully explored in the research underlying the original generalization, each circumstance thus a potential boundary condition for the "law" one seeks to apply. Caution is in order, for the history of the behavioral and social sciences abounds with examples of the simple causal generalizations of today becoming the first- and second-order interaction effects of tomorrow (Cronbach, 1975; Gergen, 1978; McGuire, 1985).

A set of predictively powerful, reliably documented generalizations is, in short, not enough; one needs some systematic way of assessing whether general principles apply to specific cases. In addition to an integrative theory, we require a diagnostic checklist for assessing whether the antecedent conditions for the activation of a given generalization are present. Such a checklist will not, moreover, be easy to devise (see Griffiths, forthcoming). The most divisive policy debates often focus on what we call from a theoretical point of view "antecedent conditions." In the post-World War II era, for instance, there has been enormous disagreement over what mixture of deterrence and reassurance is most appropriate in U.S. dealings with the Soviet Union. This debate has not hinged on the generic wisdom of a fair-but-firm strategy; it has hinged on the assessment of Soviet geopolitical intentions. Is the Soviet Union a dangerously expansionist power prepared to take large risks to achieve highly ambitious goals (R. Osgood, 1981; Wildavsky, 1983)? Or is

the Soviet Union best thought of as a conservative status quo power preoccupied with minimizing internal and external threats to its own security (White, 1984)? Or is Soviet foreign policy guided by some complex mixture of defensive and opportunistic offensive motives, with the relative importance of motives depending on the issue domain and leadership period under consideration? How one answers these questions has important implications for the emphasis one places on deterrence versus reassurance in U.S. national security policy.

Linking up theory to practice requires methods for systematically sizing up specific situations. Here the behavioral and social sciences blur into the arts of diplomacy and conflict management. This is not to say that the behavioral and social sciences have nothing to offer to the practice of international relations. These disciplines offer a variety of qualitative and quantitative techniques for predicting future trends in the behavior of nation-states (Choucri and Robinson, 1978). These disciplines also highlight the dangers of cognitive conceit (of thinking we know more than we do), point to possible correctives of judgmental biases such as overconfidence and, most crucial of all, remind us of the importance of stating our hypotheses concerning the nature of the adversary in falsifiable form (be prepared to state what would make us change our minds). But the behavioral and social sciences can apparently take us only so far. Expert observers of the Soviet Union still disagree sharply at this time over what Soviet geopolitical goals were at the time of the invasion of Afghanistan or, for that matter, over what the long-term goals of Gorbachev's foreign policy currently are.

The inferential difficulties also do not end here. Even if one had a surefire method of determining that a general principle did apply to a specific case, one would still confront the profound problem of operationalizing the theoretical advice. For example, what exactly does it mean to say that the United States should pursue a reciprocity strategy in its dealings with the Soviet Union? Reciprocity can be operationalized in a seemingly infinite variety of ways. Does it mean adopting some variant of Osgood's (1962) graduated and reciprocal initiatives in tension reduction (GRIT) proposal in which one superpower attempts to defuse tensions through a series of carefully planned and announced concessions? And what exactly should those concessions be? Should the United States announce a no-first-use policy? Should the Soviet Union have persisted with its recent unilateral nuclear test moratorium? How does one know that in operationalizing a reciprocity strategy in a particular way that one has struck the right balance between conciliatoriness and resistance to exploitation? Presumably some kind of corrective feedback mechanism needs to be built into the policy formula. The key problem then becomes calibrating one's responses to those of the other side: how does one decide

whether a given response by the other side is sufficiently conciliatory or refractory to warrant a response in kind?

Once again, the behavioral and social sciences can only take us so far. It is not possible to deduce specific policy prescriptions from these abstract bodies of knowledge (no more than it is possible to deduce a medical diagnosis of a particular patient from the biological sciences). The behavioral and social sciences do, however, highlight issues that prudent policymakers should take into account if they wish to avoid war in international confrontations. George's (George et al., 1971; George, 1980) work on the use of coercive diplomacy in the context of crisis management is an excellent illustration of work in this vein. In discussing the practice of coercive diplomacy in international politics, George did not presume to tell policymakers whether they should use force or threats of force in specific situations. On the basis of his own inductive-historical research, he did, however, identify several generic problems that policymakers need to solve if they are to be successful in a diplomatic crisis at both protecting "vital national interests" and avoiding war. For example, when considering the use of the strategy of coercive diplomacy, policymakers should ask themselves:

1. What are the risks (often considerable) of presenting an ultimatum that specifies a deadline for compliance? Can the risks be controlled?

2. How should one deal with the conflict between the need to pressure the opponent into compliance (cease attack, withdraw missiles) and the need to slow the pace of events to give the opponent time to evaluate the situation?

3. How should one calibrate the intensity and timing of threats?

4. How should threats be presented? (The linguistic, cultural, and political context can be critical determinants of whether threats backfire.) Should threats be coupled with rewards in a carrot-and-stick package that makes compliance the most attractive option? How threatening should the consequences of noncompliance be? How appealing should the consequences of compliance be? How can rewards and threats be designed to augment rather than negate each other?

These guidelines highlight the complexity of the issues and the variety of "things that can go wrong" in crisis decision making. To be sure, following these procedural guidelines is neither a necessary nor a sufficient condition for ensuring a good outcome. Policymakers may sometimes impulsively choose policies that, in hindsight, appear wise. And policymakers may sometimes carefully choose policies that, in hindsight, appear disastrous. These is no simple, all-encompassing formula for coping with the complexities of crisis management. Given what we know, however, of crisis decision making and intergroup negotiation under stress, it is reasonable to conclude that foreign-policymakers who heed these guidelines are less likely to make calamitous

miscalculations than policymakers who ignore the guidelines. George's (1980) prescriptions for crisis management do not tell us what to do, but they do tell us how to structure our thinking. The guidelines are similar in this regard to the "fault trees" that engineers use to diagnose the diverse ways in which complex physical systems can fail (Fischhoff et al., 1978).

If it is indeed impossible and perhaps undesirable to "reduce" international relations to an exercise in applied behavioral and social science, where does this leave us?

A simplistic answer is that we are left with the necessity of individual judgment. It will not be possible any time in the foreseeable future to deduce "optimal policies" (optimal in the sense of maximizing policymakers' values) from theory or research in the behavioral and social sciences. Policymakers of the future will have to rely as they do now on subjective judgment and their own often implicit crude causal theories—theories that are sometimes as inchoate as "no more Munichs" or "no more Vietnams" (Neustadt and May, 1986). A more sophisticated answer is that although we may never escape the necessity of individual judgment, we can work to ensure that the judgments of policymakers are well informed by the richer, more explicit, more differentiated—albeit sometimes fallible—knowledge base provided by the behavioral and social sciences.

It is all too easy to puncture the prescriptive pretensions of the behavioral and social sciences. The power of these disciplines to yield solutions to societal problems often seems meager compared to the power of the biological and physical sciences. And a good case can be made that excessive claims have been made in the past on behalf of the behavioral and social sciences. But if hubris is a vice to be avoided so, too, is excessive modesty. A great deal of evidence has accumulated—on social judgment and choice processes, bargaining and negotiation processes, influence processes, the functioning of individuals and organizations under stress, and the dynamics of public opinion—that should be kept in mind in public debates on key issues of international security. The appropriate benchmark of comparison is not "Has research on war and peace attained the paradigmatic consensus that prevails in certain other sciences?" but rather, "If knowledge from the behavioral and social sciences is not used to inform these debates, what types of knowledge will be used?" There is no value-neutral option here. To withhold information is as consequential an act as to release it.

The behavioral and social sciences cannot replace individual judgment, but they can sharpen, refine, and inform it. Historians and political scientists have noted that the commonsense reasoning of foreign policy elites is far from infallible. Policymakers, it has been observed, are prone to essentially the same cognitive biases and errors as ordinary mortals. They are often too quick

to draw strong conclusions from weak evidence and too slow to modify those conclusions in response to new evidence (George, 1980; Jervis, 1976; Neustadt and May, 1986). The result is often the drawing of sweeping, undifferentiated generalizations from currently salient historical precedents ("appeasement does not work" or "conventional armies cannot defeat guerrillas with strong indigenous support") and an insensitivity to differences between current problems and these popular historical precedents ("if we don't build this weapon system, we will be repeating the errors of appeasement," "if we do send troops into this country, we will be repeating the errors of Vietnam or Afghanistan"). The behavioral and social sciences have created a number of institutional checks on these types of sweeping, undifferentiated causal claims. It is incumbent on an investigator who advances such a claim to document its universal applicability and to state the claim with sufficient precision that other investigators will have a reasonably clear idea of what evidence will count either as support for or as refutation to the claim. Few claims survive this methodological screening process. Most generalizations, as is apparent from the contributions to *Behavior, Society, and Nuclear War*, have had to be qualified or circumscribed, often sharply so. Causation in international politics tends to be complex (to involve variables from a number of different levels of analysis), interactive (the effects of variables at one level of analysis often depend on the state of variables at other levels of analysis), and difficult to identify with confidence and precision (the limited number of observations, the large number of confounding variables, and the fallibility of our research methods make it difficult to disentangle competing causal hypotheses).

Behavior, Society, and Nuclear War illustrates that we have made tangible progress toward clarifying the underlying processes that affect both the likelihood of war in general and of nuclear war in particular. It also illustrates how difficult it is to make progress in this area. Readers who are looking for elegantly axiomatized theories and empirical consensus will be disappointed. Nonetheless, what has been achieved should not be minimized. We have learned a good deal on both the theoretical and the methodological fronts and, perhaps, most important, we have learned a good deal about the limits of our knowledge. Knowledge of our ignorance—especially in a policy domain where confident, even glib, causal assertions are so common—can be a major contribution in itself. The most important service the behavioral and social sciences can currently provide to the policymaking community may well be to make thoughtful skepticism respectable: to sensitize those who make key decisions to the uncertainty surrounding our understanding of international conflict and to the numerous qualifications that now need to be attached to simple causal theories concerning the origins of war.

In the same spirit, *Behavior, Society, and Nuclear War* should not be viewed as a final or definitive statement of our knowledge, but as a product of an ongoing process of inquiry. The process of adapting behavioral and social science methods to the study of international security issues is a massive undertaking that requires careful analysis of the strengths and weaknesses of different methods and, for many problems, the invention of hybrid approaches that combine the distinctive strengths of these methods (e.g., comparative case studies combined with game theoretic modeling or laboratory simulations). The process of formulating tentative theoretical generalizations and subjecting them to criticism and revision is an equally massive undertaking that requires careful and patient empirical evaluations of a confusing variety of claims, counterclaims, and revised claims. The volumes of *Behavior, Society, and Nuclear War* represent the initial stages of these efforts. Our goal has been to bring together those methodological and theoretical advances that, in our collective judgment, promise the most fruitful payoffs in the next stages of this ongoing process.

Notes

1. It is worth noting that although scholars typically study processes at only one level of analysis (the departmental structure of research universities encourages such specialization), policymakers do not have this analytical luxury. They must deal with complex, real-world events that defy neat theoretical taxonomies. Policymakers need to have at least a crude intuitive understanding of when processes at one or another level of analysis have become particularly critical determinants of policy outcomes. They need to make judgments such as: Has the leadership of nation x misjudged our likely reaction to this policy? Is the leadership pursuing goals different from those that we initially supposed? How much latitude for maneuver does the leadership of nation x have given the domestic political and economic constraints within which it must work? Does this response of nation x reflect the intentions of the central leadership, or is it an act initiated by a particular influence group or bureaucracy within the government? One need only think of the vigorous policy debates within the United States that have been triggered by events such as the Cuban missile crisis, the Third World debt crises, and the Korean Airlines flight 007 disaster to recognize the enormous impact that level-of-analysis attributions can have on policy preferences.

2. Levels of analysis represent broad conceptual frameworks within which it is

possible to articulate and refine numerous testable "theories of the middle range" (Merton, 1957). Thus, there is no single psychological theory of individual decision making, no single organizational theory of bureaucratic politics, and no single systemic theory of how the balance of power affects the likelihood of war. This internal complexity of levels of analysis makes it extremely difficult, perhaps at present impossible, to falsify the claim that a particular level of analysis is necessary or sufficient for explaining a particular phenomenon. As soon as one has rejected hypotheses derived from one middle-range theory, another middle-range theory emerges to replace it. Presumably, a limit must be placed on this process; repeated failures to reconcile evidence with middle-range theories from a given level of analysis are reminiscent of what Lakatos (1970) described as "degenerative" research programs. When more intellectual energy goes into thinking of post hoc interpretations to defend existing theory than goes into thinking of ways of extending existing theory to new evidence, the time has probably come to reevaluate the viability of the entire research program.

3. We need to be careful not to assume that all role-induced attitude shifts are purely opportunistic. Changes in roles may expose policymakers to new evidence and, in turn, produce genuine shifts in intellectual perspective. It is difficult in analysis that, in turn, produce genuine shifts in intellectual perspective. It is difficult in any given case to disentangle opportunistic from information-driven attitude change.

4. Domestic political pressures do not, however, operate in only this direction. Historians have documented many cases in which, were it not for influential antiwar domestic constituencies, national leaders almost certainly would have pursued more bellicose policies (for example, Roosevelt prior to 1941 or the Johnson administration during the Vietnam War).

5. Advocates of quantitative, variable-centered research are not, of course, without counterarguments. If nomothetic researchers were simply measuring radically different properties of behavior in different situations and arbitrarily categorizing those properties under the same variable label, one would not expect statistically powerful or replicable relationships to emerge from studies conducted by these researchers. Since nomothetic research sometimes reveals powerful and replicable relationships, the critique is overstated. Advocates of qualitative idiographic approaches can still, however, claim that quantitative researchers typically treat context-specific meanings as statistical noise or error variance and that, as a result, seriously oversimplify reality. Phenomena that represent artifacts or nuisance variables from a nomothetic point of view may be of central interest from an idiographic perspective.

6. Although the tit-for-tat strategy accumulated large numbers of points against other response programs in Axelrod's computer simulation tournament, that does not mean tit-for-tat is the key to survival in the international environment (Jervis, 1988). One can raise a variety of objections to the strategy. From a "dovish" perspective, tit-for-tat may be too tough. Once one has entered into a competitive response cycle, it is unclear how tit-for-tat can get one out. Someone has to take the conciliatory initiative (see Osgood's [1962] discussion of GRIT; Larsen, 1987). From a "hawkish" perspective, tit-for-tat may be too soft. As Axelrod (1984) notes, tit-for-tat never actually won any individual game in the computer tournament. If the primary goal in international politics is defined as maximizing one's *relative* gain, a response strategy that can at

best tie loses much of its attractiveness. Finally, from a psychological point of view, tit-for-tat may simply not work very well in a world in which perceptual errors occur—in which decision makers frequently misclassify cooperative behavior as competitive and competitive behavior as cooperative (Are Gorbachevian arms control concessions motivated by the desire to reach a stable, mutually beneficial modus vivendi with the West? Or are the concessions attempts to gain breathing time for a political-economic system that, once recovered, will pose an even more severe threat to Western security?). From this analytic perspective key questions become "How high an error rate can tit-for-tat withstand?" and "How high is the actual error rate in international politics?" In brief, the game-theoretic analysis ultimately has to be grounded in psychological and political reality.

7. Even this pretty robust generalization requires qualification. For instance, Nazi Germany in 1939 and Imperial Japan in 1941 appear to have been willing to take very large risks to expand national power. These observations can, however, be readily reconciled with some form of subjective expected utility theory. Germany, it could be argued, sought to recover from the enormous losses of World War I (hence its willingness to take risks), and Japan, it could be argued, feared that the military-economic balance of power would shift progressively against it unless decisive action were taken (Russett, 1967).

8. Work on framing effects on decision making (Tversky and Kahneman, 1981) suggests that public support for arms control proposals that require complex trade-offs is likely to be volatile and to depend very much on how salient the question makes the "loss" and "gain" sides of the trade-off equation. This perspective also suggests that in a multidimensional, asymmetric strategic environment (in which the two sides possess distinctive and difficult-to-compare strengths and weaknesses [Steinbruner, 1985]), opponents of arms control will *ceteris paribus* have a built-in psychological edge in the battle for public opinion as a result of the tendency for losses to loom larger than gains.

References

Abelson, R.P., and A. Levi. 1985. Decision-making and decision theory. In E. Aronson and G. Lindzey, eds., *Handbook of Social Psychology*, Vol. 2. Hillsdale, N.J.: Erlbaum.

Allison, G.T. 1971. *Essence of Decision*. Boston: Little, Brown.

Allison, G.T., A. Carnesale, and J.S. Nye. 1985. *Hawks, Doves, and Owls: An Agenda for Avoiding Nuclear War*. New York: Norton.

Almond, G., and T. Genco. 1977. Clouds, clocks, and the study of politics. *World Politics* 29:489–522.

Art, R.J. Forthcoming. U.S. nuclear policy and the political process. In P.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Strn, and C. Tilly, eds., *Behavior, Society, and Nuclear War*. Vol. 3. New York: Oxford University Press.

Axelrod, R. 1976. *Structure of Decision*. Princeton, N.J.: Princeton University Press.

———. 1984. *The Evolution of Cooperation*. New York: Basic Books.

Axelrod, R., and R.O. Keohane. 1985. Achieving cooperation under anarchy: Strategies and institutions. *World Politics* 38:226–254.

Blair, B.G. 1987. Alerting in crisis and conventional war. In A.B. Carter, J.D. Steinbruner, and C.A. Zracket, eds., *Managing Nuclear Operations*. Washington, D.C.: The Brookings Institution.

Bracken, P. 1983. *Command and Control of Nuclear Forces*. New Haven, Conn.: Yale University Press.

Brams, S. 1985. *Superpower Games*. New Haven, Conn.: Yale University Press.

Brewer, M., and R.M. Kramer. 1985. The psychology of intergroup relations and behavior. *Annual Review of Psychology*, 36:219–244.

Bueno de Mesquita, B. 1981. *The War Trap*. New Haven, Conn.: Yale University Press.

———. 1985. The war trap revisited: A revised expected utility model. *American Political Science Review* 79:157–176.

Burnstein, E., and M.L. Berbaum. 1983. Stages in group decision-making: The decomposition of historical narratives. *Political Psychology* 4:531–561.

Campbell, D.T., and D.W. Fiske. 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 56:81–105.

Chertkoff, J.M., and S.L. Baird. 1971. Applicability of the big lie technique and the last clear chance doctrine to bargaining. *Journal of Personality and Social Psychology* 20:298–303.

Chertkoff, J.M., and J.K. Esser. 1976. A review of experiments in explicit bargaining. *Journal of Experimental Social Psychology* 12:222–236.

Choucri, N., and R.C. North. 1975. *Nations in Conflict*. San Francisco, Calif.: Freeman.

Choucri, N., and T.W. Robinson. 1978. *Forecasting in International Relations*. San Francisco, Calif.: Freeman.

Converse, P.E. 1964. The nature of belief systems in mass publics. In D.E. Apter, ed., *Ideology and Discontent*. New York: Free Press.

Coser, L. 1956. *The Function of Social Conflict*. New York: Free Press.

Cronbach, L.J. 1975. Beyond the two disciplines in scientific psychology. *American Psychologist* 30:116–127.

Dyson, F. 1984. *Weapons and Hope*. New York: Harper & Row.

Eckstein, H. 1975. Case studies in political science. In N. Polsby and F.I. Greenstein, eds., *Handbook of Political Science*, Vol. 6. Reading, Mass.: Addison-Wesley.

Einhorn, H., and R.M. Hogarth. 1981. Behavioral decision theory. *Annual Review of Psychology* 32:53–88.

Etheredge, L.S. 1978. *A World of Men*. Cambridge, Mass.: MIT Press.

Fischhoff, B. Forthcoming. Nuclear decisions: Cognitive limits to the thinkable. In P.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Stern, and C. Tilly, eds., *Behavior, Society, and Nuclear War*, Vol. 2. New York: Oxford University Press.

Fischhoff, B., P. Slovic, and S. Lichtenstein. 1978. Fault trees: Sensitivity of estimated failure estimates to problem representation. *Journal of Experimental Psychology: Human Perception and Performance* 2:330–344.

Gallie, W.B. 1968. *Philosophy and the Historical Understanding*, 2d. ed. New York: Schocken.

Geertz, C. 1973. *The Interpretation of Culture*. Princeton, N.J.: Princeton University Press.

George, A.L. 1959. *Propaganda Analysis*. White Plains, N.Y.: Row, Peterson.

——. 1979. Case studies and theory development. In P.G. Laure ed., *Diplomatic History: New Approaches*. New York: Free Press.

——. 1980. *Presidential Decision-Making in Foreign Policy: The Effective Use of Information and Advice*. Boulder, Col.: Westview.

George, A.L., P.J. Farley, and A. Dallin. 1988. *U.S.-Soviet Security Cooperation: Achievements, Failures, Lessons*. Oxford, England: Oxford University Press.

George, A.L., D. Hall, and W. Simons. 1971. *The Limits of Coercive Diplomacy*. Boston: Little, Brown.

George, A.L., and T.J. McKeown. 1985. Case studies and theories of organizational decision making. In *Advances in Information Processing*, Vol. 2. Greenwich, Conn.: JAI Press.

George, A.L., and R. Smoke. 1974. *Deterrence in American Foreign Policy: Theory and Practice*. New York: Columbia University Press.

Gergen, K.J. 1978. Experimentation in social psychology: A reappraisal. *European Journal of Social Psychology* 8:507–527.

Greenstein, F.I. 1975. Personality and politics. In N. Polsby and F.I. Greenstein, eds., *Handbook of Political Science*, Vol. 2. Reading, Mass.: Addison-Wesley.

Griffiths, F. Forthcoming. Mutual perceptions of intentions among Soviet and U.S. policy elites. In R.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Stern, and C. Tilly, eds., *Behavior, Society, and Nuclear War*, Vol. 3. New York: Oxford University Press.

Halperin, M. 1974. *Bureaucratic Politics and Foreign Policy*. Washington, D.C.: The Brookings Institution.

Hempel, C.G. 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Hoffman, S. 1960. *Contemporary Theory in International Relations*. Englewood Cliffs, N.J.: Prentice-Hall.

Holsti, O.R. 1972. *Crisis Escalation War*. Montreal, Canada: McGill Queens University Press.

——. 1976. Foreign policy formation viewed cognitively. In R. Axelrod, ed., *Structure of Decision*. Princeton, N.J.: Princeton University Press.

Holsti, O.R., R.A. Brody, and R.C. North. 1969. The management of international crisis: Affect and action in American-Soviet relations. In D.G. Pruitt and R.C. Snyder, eds., *Theory and Research on the Causes of War*. Englewood Cliffs, N.J.: Prentice-Hall.

Institute of Medicine. 1986. *The Medical Implications of Nuclear War*. Washington, D.C.: National Academy Press.

Intriligator, M.D., and D.L. Brito. 1984. Can arms races lead to the outbreak of war? *Journal of Conflict Resolution* 28:63–84.

Janis, I.L. 1982. *Victims of Groupthink*, 2nd ed. Boston: Houghton-Mifflin.

Janis, I.L., and L. Mann. 1977. *Decision Making*. New York: Free Press.

Jervis, R. 1969. The costs of the quantitative study of international relations. In K. Knorr and J. Rosenau, eds., *Contending Approaches in International Politics*. Princeton, N.J.: Princeton University Press.

——. 1970. *The Logic of Images in International Relations*. Princeton, N.J.: Princeton University Press.

——. 1976. *Perception and Misperception in International Politics*. Princeton, N.J.: Princeton University Press.

——. 1978. Cooperation under the security dilemma. *World Politics* 30:167–214.

——. 1988. Realism, game theory and cooperation. *World Politics* 41:317–349.

Kahneman, D., P. Slovic, and A. Tversky, eds. 1982. *Judgment Under Uncertainty*. Cambridge, England: Cambridge University Press.

Kahneman, D., and A. Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica* 47:263–291.

Kaplan, A. 1964. *The Conduct of Inquiry*. Scranton, Penn.: Chandler.

Kaplan, F. 1982. *Wizards of Armageddon*. New York: Simon and Schuster.

Katz, A.M. 1982. *Life after Nuclear War: The Economic and Social Impacts of Nuclear Attacks on the United States*. Cambridge, Mass.: Ballinger.

Kelman, H., and A. Bloom. 1973. Assumptive frameworks in international politics. In J. Knutso, ed., *Handbook of Political Psychology*. San Francisco, Calif.: Jossey-Bass.

Kennedy, P. 1987. *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*. New York: Random House.

Lakatos, I. 1970. Falsification and the methodology of scientific research programs. In I. Lakatos and A. Musgrave, eds., *Criticism and Growth of Knowledge*. London: Cambridge University Press.

Larsen, D. 1987. Game theory and the psychology of reciprocity. Unpublished manuscript, Columbia University.

Lebow, R.N. 1981. *Between Peace and War*. Baltimore, Md.: Johns Hopkins University Press.

——. 1987. Deterrence failure revisited. *International Security* 12(1):197–213.

Leng, R.J. 1983. When will they ever learn: Coercive bargaining in recurrent crises. *Journal of Conflict Resolution* 27:379–420.

Leng, R.J., and H.G. Wheeler. 1979. Influence strategies, success, and war. *Journal of Conflict Resolution* 23:655–684.

Majeski, S.J., and D.J. Sylvan. 1984. Simple choices and complex calculations. *Journal of Conflict Resolution* 38:316–340.

Maoz, Z. 1981. The decision to raid Entebbe. *Journal of Conflict Resolution* 25:677–707.

March, J.G. 1978. Bounded rationality, ambiguity, and the engineering of choice. *Bell Journal of Economics* 9:587–603.

McClosky, H. 1967. Personality and attitude correlates of foreign policy orientation. In J.N. Rosenau, ed., *Domestic Sources of Foreign Policy*. New York: Free Press.

McGuire, W.J. 1985. The nature of attitudes and attitude change. In G. Lindzey and E. Aronson, eds., *Handbook of Social Psychology*, 3rd ed. Reading, Mass.: Addison-Wesley.

Melman, S. 1970. *Pentagon Capitalism: The Political Economy of War*. New York: Oxford University Press.

Merton, R.K. 1957. *Social Theory and Social Structure*. Glencoe, Ill.: Free Press.

Modelski, G. 1987. *Long Cycles in World Politics*. Seattle, Wash.: University of Washington Press.

Mueller, J.E. 1973. *War, Presidents, and Public Opinion*. New York: Wiley.

Neustadt, R.E., and E.R. May. 1986. *Thinking in Time: The Uses of History for Decision Makers*. New York: Free Press.

Nisbett, R.E., and L. Ross. 1980. *Human Inference*. Englewood Cliffs, N.J.: Prentice Hall.

Organski, A.F.D., and J. Kugler. 1980. *The War Ledger*. Chicago: University of Chicago Press.

Orme, J. 1987. Deterrence failures: A second look. *International Security* 11(3):3–40.

Osgood, C.E. 1962. *An Alternative to War or Surrender*. Urbana, Ill.: University of Illinois Press.

Osgood, R.E. 1981. *Containment, Soviet Behavior and Grand Strategy*. Berkeley, Calif.: Institute of International Studies.

Oye, K.A. 1986. *Cooperation Under Anarchy*. Princeton, N.J.: Princeton University Press.

Pfeffer, J. 1985. Organizations and organization theory. In E. Aronson and G. Lindzey, eds., *Handbook of Social Psychology*, Vol. 1. Hillsdale, N.J.: Erlbaum.

Pruitt, D.G., and S.A. Lewis. 1975. Development of integrative solutions in bilateral negotiation. *Journal of Personality and Social Psychology* 31:621–633.

Pruitt, D.G., and J. Rubin. 1986. *Social Conflict: Escalation, Stalemate and Settlement*. New York: Random House.

Richardson, L.F. 1960. *Arms and Insecurity*. London: Stevens and Sons Ltd.

Rosecrance, R. 1963. *Action and Reaction in World Politics*. Boston: Little, Brown.

Rosenau, J. 1970. Theories and pre-theories of foreign policy. In R.B. Farrell, ed., *Approaches to Comparative and International Politics*. Evanston, Ill.: Northwestern University Press.

Russett, B.M. 1967. Pearl Harbor: Deterrence theory and decision theory. *Journal of Peace Research* 4:89–106.

Sagan, C. 1983. Nuclear war and climatic catastrophe. *Foreign Affairs* 62:257–292.

Sarkesian, S.C., ed. 1972. *The Military-Industrial Complex: A Reassessment*. Beverly Hills, Calif.: Sage.

Schell, J. 1980. *The Fate of the Earth*. New York: Knopf.

Schelling, T.C. 1966. *Arms and Influence*. New Haven, Conn.: Yale University Press.

Schlenker, B. 1985. *The Self in Social Life*. New York: Academic Press.

Schumann, H., and S. Presser. 1981. *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*. New York: Academic Press.

Shubik, M. Forthcoming. Models of strategic behavior and nuclear deterrence. In P.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Stern, and C. Tilly, eds., *Behavior, Society, and Nuclear War*. Vol. 2. New York: Oxford University Press.

Simon, H.A. 1985. Human nature in politics: The dialogue of psychology with political science. *American Political Science Review* 79:293–304.

Singer, D. 1972. *The Scientific Study of Politics: An Approach to Foreign Policy Analysis*. New York: General Learning Press.

Singer, J.D. 1980. Accounting for international war: The state of the discipline. *Annual Review of Sociology* 6:349–367.

Snyder, G.H., and P. Diesing. 1977. *Conflict Among Nations: Bargaining Decision Making, and System Structure in International Crises*. Princeton, N.J.: Princeton University Press.

Stein, J.G.. Forthcoming. Deterrence and reassurance. In P.E. Tetlock, J.L. Husbands, R. Jervis, P.C. Stern, and C. Tilly, eds., *Behavior, Society, and Nuclear War*, Vol. 2. New York: Oxford University Press.

Steinbruner, J.D. 1974. *A Cybernetic Theory of Decision*. Princeton, N.J.: Princeton University Press.

———. 1985. U.S. and Soviet security perspectives. *Bulletin of the Atomic Scientists* 22:89–93.

———. 1987. Choices and trade-offs. In A.B. Carter, J.D. Steinbruner, and C.A. Zracket, eds., *Managing Nuclear Operations*. Washington, D.C.: The Brookings Institution.

Stephenson, G.M., B.H. Kniveton, and I.E. Morely. 1977. Interaction analysis of an industrial wage negotiation. *Journal of Occupational Psychology* 50:231–241.

Streufert, S., and S. Streufert. 1978. *Behavior in the Complex Environment*. Washington, D.C.: Winston & Sons.

Suedfeld, P., and P.E. Tetlock. 1977. Integrative complexity of communications in international crises. *Journal of Conflict Resolution* 21:169–184.

Talbott, S. 1985. *Deadly Gambits*. New York: Knopf.

Tetlock, P.E. 1979. Identifying victims of groupthink from public statements of decision-makers. *Journal of Personality and Social Psychology* 37:1314–1324.

———. 1985a. Integrative complexity of American and Soviet foreign policy statements: A time-series analysis. *Journal of Personality and Social Psychology* 49:1565–1585.

———. 1985b. Accountability: The neglected social context of judgment and choice. In B.M. Staw and L. Cummings, eds., *Research in Organizational Behavior*, Vol. 1. Greenwich, Conn.: JAI Press.

Tetlock, P.E., and J.I. Kim. 1987. Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology* 52:700–709.

Tetlock, P.E., and C. McGuire. 1986. Cognitive perspectives on foreign policy. In S. Long ed., *Political Behavior Annual*, Vol. 1. Boulder, Col.: Westview. Reprinted in R. White, ed., 1986. *Psychology and the Prevention of Nuclear War*. New York: New York University Press.

Theis, W. 1984. *When Governments Collide*. Berkeley, Calif.: University of California Press.

Tversky, A., and D. Kahneman. 1981. The framing of decisions and the rationality of choice. *Science* 221:453–458.

Verba, S. 1961. Assumptions of rationality and non-rationality in the international system. In K. Knorr and S. Verba, eds., *The International System: Theoretical Essays*. Princeton, N.J.: Princeton University Press.

———. 1967. Some dilemmas in comparative research. *World Politics* 20:111–127.

Walcott, C., P.T. Hopmann, and T.D. King. 1977. The role of debate in negotiation. In D. Druckman, ed., *Negotiations: Social Psychological Perspectives*. Beverly Hills, Calif.: Sage.

Walker, S. 1977. The interface between beliefs and behavior: Henry Kissinger's operational code and the Vietnam War. *Journal of Conflict Resolution* 21:129–168.

Walsh, W.H. 1967. *Philosophy of History: An Introduction*. New York: Harper & Row.

Waltz, K.N. 1959. *Man, the State, and War*. New York: Columbia University Press.

———. 1979. *Theory of International Politics*. Reading, Mass.: Addison Wesley.

Weick, K.E. 1979. Cognitive processes in organizations. In B.M. Straw, ed., *Research in Organizational Behavior*. Greenwich, Conn.: JAI Press.

White, R. 1984. *Fearful Warriors: A Psychological Study of U.S.-Soviet Relations*. New York: Free Press.

Wiegele, T.C. 1979. Signal leakage and the remote psychological assessment of foreign policy elites. In L. Falkowski, ed., *Psychological Models in International Politics*. Boulder, Col.: Westview.

Wildavsky, A. 1983. *Beyond Containment: Alternative American Policies toward the Soviet Union*. San Francisco: Institute of Contemporary Studies.

Wilson, W. 1971. Reciprocation and other techniques for inducing cooperation in the prisoner's dilemma game. *Journal of Conflict Resolution* 15:167–195.

Yarmolinsky, A. 1970. *The Military Establishment*. New York: Harper & Row.

Contributors and Editors

DANIEL DRUCKMAN is a study director at the National Research Council and adjunct professor of conflict management at George Mason University. He has been a senior scientist at Booz Allen & Hamilton and the Mathtech Scientist at Mathematica, Inc. His primary interests are in the ares of interparty conflict resolution, international negotiations, nonverbal communication, political analysis, and modeling methodologies, including simulation. His publications include *Negotiations: Social-Psychological Perspectives* (Sage, 1977), *Nonverbal Communication: Survey, Theory, and Research* (Sage, 1982), and *Political Stability in the Philippines: Framework and Analysis* (University of Denver, 1986). Druckman received a Ph.D. in social psychology from Northwestern University.

OLE HOLSTI is a professor of political science at Duke University. His research interests include foreign policy, crisis decision making, public opinion and foreign policy, international relations, and international relations theory. He is author of *Crisis, Escalation, War* (McGill Queens University Press, 1972) and coauthor of *American Leadership in World Affairs: UN and the Breakdown of Consensus* (with James N. Rosenau; Allen & Unwin, 1984). Holsti received a B.A. from Stanford, an M.A.T. from Wesleyan University, and a Ph.D. in political science from Stanford University.

P. TERRENCE HOPMANN is a professor of political science and director of the Institute of International Studies at Brown University. His primary interests are in the areas of international negotiation, arms control, and the history of U.S.-Soviet arms negotiations. His publications include *Rethinking the Nuclear Weapons Dilemma in Europe* (Macmillan, 1988) and *Unity and Disintegration in International Alliances* (Wiley, 1973). He holds an A.B. from Princeton University and a Ph.D. in political science from Stanford University.