

Comparative Political Studies

<http://cps.sagepub.com>

Case Study Methods in the International Relations Subfield

Andrew Bennett and Colin Elman
Comparative Political Studies 2007; 40; 170
DOI: 10.1177/0010414006296346

The online version of this article can be found at:
<http://cps.sagepub.com/cgi/content/abstract/40/2/170>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Comparative Political Studies* can be found at:

Email Alerts: <http://cps.sagepub.com/cgi/alerts>

Subscriptions: <http://cps.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 20 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://cps.sagepub.com/cgi/content/abstract/40/2/170#BIBL>

Case Study Methods in the International Relations Subfield

Andrew Bennett

Georgetown University, Washington, D.C.

Colin Elman

Arizona State University, Tempe

This article reviews the key role that case study methods have played in the study of international relations (IR) in the United States. Case studies in the IR subfield are not the unconnected, atheoretical, and idiographic studies that their critics decry. IR case studies follow an increasingly standardized and rigorous set of prescriptions and have, together with statistical and formal work, contributed to cumulatively improving understandings of world politics. The article discusses and reviews examples of case selection criteria (including least likely, least and most similar, and deviant cases); conceptual innovation; typological theories, explanatory typologies, qualitative comparative analysis, and fuzzy-set analysis; process tracing; and the integration of multiple methods.

Keywords: *qualitative methods; case selection; conceptual innovation; explanatory typologies; multimethod analysis; process tracing*

Qualitative methods have for many decades played a key role in the study of international relations (IR) in the United States. In sharp contrast to the analysis of American politics, they have been at least as important to the IR subfield as quantitative approaches. Although we cannot cover all of such a wide-ranging literature in a single article, we do describe several prominent aspects of case-based research in the IR subfield. In addition, although the examples we cite are mainly drawn from our own specializations in security studies and foreign policy, we believe that these reflect the situation in the IR subfield as a whole. For example, John Odell (2004) argues that “case study methods have dominated the study of IPE [international political economy] over the last three decades” (p. 56).¹

Authors' Note: A previous version of this article was presented at the 2005 Annual Meetings of the American Political Science Association, Washington, D.C., September 1 to 4. For helpful comments, the authors thank James Caporaso and David Waldner.

As we noted in the essay introducing this symposium, there are many different reasons why social science subfields adopt particular approaches to analysis, including the interposition of significant events, academic entrepreneurship, and others. We believe that one important reason why qualitative methods have been important in IR research is that case study methods, especially the combination of process tracing and typological theorizing, have considerable advantages in studying complex phenomena (Bennett & Elman, 2006a; Mahoney & Goertz, 2006). Some of the most prominent subjects of study, for example wars and the formation of new international security and economic systems, often involve interaction effects among many structural and agent-based variables, path dependencies, and strategic interaction among large numbers of actors across multiple levels of analysis with private information and strong incentives to bluff or deceive other actors. In addition, although system-altering wars and economic changes are relatively infrequent phenomena, the structure of the international system is nonetheless more anarchical and subject to continual renegotiation and periodic violent challenges than the realm of domestic politics. System-altering revolutions of course happen in domestic politics as well, but in domestic politics actors make much greater investments in institutions that structure political life and reduce many of its uncertainties. The prominence of qualitative methods in IR thus reflects these methods' advantages in studying complex and relatively unstructured and infrequent phenomena that lie at the heart of the subfield.

To be sure, statistical methods and formal models have made seminal contributions in the study of IR as well, notably in validating the importance of power differentials, establishing the (still contested) proposition that democracies rarely if ever fight wars against one another, and providing insight into the problems of strategic interaction, imperfect information, and credible commitments. The study of IR will undoubtedly progress most rapidly through a combination of methodological approaches. Nevertheless, formal and statistical methods have not been used in as great a proportion of research in IR as they have in the American politics subfield (and, to a lesser extent, in comparative politics). We believe that is in large part because of the complexity of IR and the ubiquity of phenomena that in many respects are *sui generis*, thus rendering many puzzles in IR difficult to model formally and to test statistically.

Almost every major research program in the IR subfield has benefited from the application of case study methods.² In this article, we discuss just a few of these many excellent examples of qualitative research. Although it has been ubiquitous, qualitative research in IR has not been without its critics.³

One of the biggest challenges for case study approaches has been to demonstrate that scholars have been using a commonly held collection of methods, thus producing a progressively improving set of understandings about war, peace, and other international phenomena. Skeptics have been unconvinced, arguing instead that much of the research amounts to a collection of unconnected, atheoretical, and idiographic case studies. Zeev Maoz (2002), for example, suggests that “one often gets the impression that the use of the case study absolves the author from any kind of methodological considerations. Case studies have become in many cases a synonym for freeform research where everything goes” (pp. 164-165). We argue that although Maoz’s assessment of contemporary literature in the IR subfield may apply to some case studies carried out with insufficient methodological rigor, and although case study researchers need to continue efforts to make their work more replicable, his critique is becoming increasingly inapt. The IR subfield includes several outstanding case studies that have contributed, together with statistical and formal work, to cumulatively improving understandings of world politics.

We analyze below examples of IR research using qualitative methods, including studies of deviant cases, least-likely cases, and most-similar cases, and examples of conceptual innovation, typological theorizing, refinements of measurement and case codings, and process tracing.⁴ We focus on exemplary works, though we at times offer some constructive critiques. We emphasize the advantages of multimethod research, whether carried out within a single study combining qualitative, statistical, and formal methods or sequentially through studies that build on earlier research using different methods. We conclude by highlighting three areas where continuing improvements in qualitative methods would be helpful: further integrating qualitative methods with formal and statistical methods, making qualitative studies more replicable by providing Internet access to primary source materials, and improving techniques for reliably assessing the perceptions and identities that are of interest to constructivist and others scholars of IR.

Studying the Right Cases: Selection Criteria for Case Study Methods

To ensure cumulation, it is vital that cases be selected self-consciously and with a view to maximizing inferential leverage. Although the selection criteria that are applicable to large-*N* studies are not always useful for qualitative approaches (D. Collier & Mahoney, 1996; Mahoney & Goertz, 2006;

cf. Geddes, 2003; King, Keohane, & Verba, 1994; Maoz, 2002, p. 169), there are other rationales for studying particular kinds of cases. Much of the strongest qualitative work in IR pays close attention to the reasons for studying cases.

“Least-Likely” Case Studies in IR

The least-likely case study relies on what Jack Levy (2002, p. 144) has labeled the “Sinatra inference”: If the theory can make it here, it can make it anywhere. Least-likely cases follow a Bayesian logic: The more surprising an outcome is relative to extant theories, the more we increase our confidence in the theory or theories that are consistent with that outcome. In the strongest instance of such logic, if a theory of interest predicts one outcome in a case, if the variables of that theory are not at extreme levels that strongly push toward that outcome, and if all of the alternative hypotheses predict a different outcome in that case, this is a least-likely case for the theory of interest. There are few examples of a theory surviving such a perfectly least-likely case, but the Bayesian logic underlying this kind of inference applies as well to a case that, although not difficult on every dimension, is nevertheless hard enough to constitute a tough test.

Matthew Evangelista (1999) provides an excellent example of a tough test case in his book on transnational actors (TNAs) and U.S.-Soviet defense and arms control policies in the cold war. Evangelista notes that cold war security issues are a “hard case” for theories that emphasize TNAs, as the realist theories that have been predominant in the study of security issues argue that states are the key actors on such issues (p. 6). Through extraordinarily detailed and tenacious process tracing, including both archival research and interviews, Evangelista demonstrates that transnational contacts, particularly those between U.S. and Soviet scientists, did indeed affect the course of U.S. and Soviet defense and arms-control policies. The strongest instance of this concerns the turnaround in Soviet policies on ballistic missile defenses. Evangelista documents that initially, when American scientists argued that building imperfect missile defenses might be counterproductive because they could increase the incentives for a first strike, the leader of a Soviet delegation of scientists found the proposal to limit defensive weapons so counterintuitive that he asked if the interpreter had correctly translated his American colleague’s remarks (p. 133). Eventually, as Evangelista demonstrates in detail, Soviet scientists became convinced by their American counterparts that spending on missile defenses would be wasteful and potentially even destabilizing, and they

successfully lobbied Soviet political leaders to embrace the 1972 SALT I Treaty limiting the testing and deployment of defenses against ballistic missiles. Evangelista gives careful attention throughout his book to alternative economic, military, and political explanations, and he concludes that TNAs had weaker effects on negotiations toward a nuclear test ban and on Soviet conventional force levels than on ballistic missile defenses.

Similarly, Peter Katzenstein (1996, p. 11) notes that the essays in his edited volume, *The Culture of National Security: Norms and Identity in World Politics*, apply constructivist theories to the hard cases of security relationships, demonstrating that even the interstate relations that are seemingly most powerfully driven by material power balances in fact are to some degree socially constructed.

Both Katzenstein's and Evangelista's volumes, however, could be more explicit and precise on the conditions under which they expect TNAs or socially constructed identities to be the most and least important determinants of security or other policies. Cases are seldom least likely for a theory in every single respect, and theories seldom hold true in contexts that are totally adverse. In the Katzenstein study, national security issues are both highly material and highly symbolic; other issue areas, such as international sporting competitions, are highly symbolic but have little effect on the security or economic well-being of states, whereas other issues might be economically consequential but carry little symbolism, and still other issues might be low in both material consequences and symbolic import. In Evangelista's study, the specialized knowledge and governmental or quasi-governmental status of the scientists involved in transnational arms control discussions gave them far more legitimacy and credibility than nonspecialist, nongovernmental actors would have had. Evangelista does note that state-level variables affect the roles of TNAs. In the strong Soviet state, it was difficult for TNAs to gain access to top leaders, but if they did so, they could powerfully affect policy, whereas in the weaker Russian state, TNAs have to contend with other state and societal actors both to be heard and to affect the implementation of policies. To develop TNA theory more cumulatively, it is necessary to build on this and other insights toward a more integrated theory on how variations in the kind of TNA involved, the issue area in question, and the characteristics of governments and their relations with one another all interact to shape the influence of TNAs.

Most-Similar and Least-Similar Case Comparisons in IR

In a most-similar case comparison, which builds on the logic of Mill's "method of difference," the key challenge is to find cases that are as similar

as possible in all but one independent variable and that differ in their outcomes. Because it is never possible to find perfectly matched cases, which would require that the cases would be matched vis-à-vis both known and unknown rival hypotheses, the second challenge is to demonstrate that the difference in the value of the independent variable of interest between the two cases, rather than the residual differences between the two cases identified by rival hypotheses, accounts for the difference in outcomes.⁵ A key method for undertaking this latter task is to use process tracing to show that the independent variable of interest that differs between the cases does in fact affect their outcomes and to show that residual differences between the cases are not connected to the difference in outcomes by any plausible causal process.

One excellent example of this technique from the democratic peace research program is James Lee Ray's (1995) most-similar case comparison of the Fashoda Crisis, a diplomatic crisis between two democracies (Britain and France) at the end of the 19th century, and the Spanish-American War, a conflict between a democracy and an autocratic state. To ensure that residual differences between the cases do not account for their differing outcomes, Ray addresses several standard kinds of confounding variables and processes, including "regression effects" (selection of cases based on extreme scores), "mortality" (the differential loss of respondents from a study, in this instance, the possibility that states may become more authoritarian as they see a war coming), and selection bias. Ray circumvents other common sources of incomparability, such as the effects of history, maturation, and changes in instrumentation, by selecting cases from the same year. Ray then addresses six other variables that might account for the different outcomes of the two cases, including proximity, power ratios, alliances, levels of economic development, militarization, and political stability. Ray's careful attention to a comprehensive list of actual and potential differences between the cases, and his use of process-tracing evidence (though not as comprehensive or detailed as that in some other studies of the Fashoda crisis), bolster his conclusions that Spain's autocracy contributed to the Spanish-American war and the democratic nature of the French and British political systems helped peacefully resolve the Fashoda crisis.

In a least-similar cases design, the researcher selects cases that are dissimilar in all but one independent variable but that share the same dependent variable. This can provide evidence that the single common independent variable helps account for the common dependent variable. Here, the task is to use process tracing to show that the common independent variable is related to the outcome through a plausible causal path and that other residual similarities between the cases are not. An example of this research design that

uses both statistical and case study methods is a study by Carol Ember, Melvin Ember, and Bruce Russett (1992) that tests whether the democratic peace finding, developed from research on modern states, applies as well to preindustrial societies. These authors conclude that despite the many differences between modern and preindustrial societies, preindustrial societies with participatory decision processes were also unlikely to fight one another.

Combining Cross-Case and Over-Time Comparisons

One research design that can generate considerable inferential leverage from the study of a few cases is the combination of cross-case and over-time (or before-after) case comparisons. This allows each case to be potentially compared in two different ways: The before of case *A* at T^0 can be compared to the after of case *A* at T^1 , whereas case *A* in one or both periods can also be compared to another case *B*, which might also be divided into two periods. Stephen Walt's (1996) *Revolution and War*, for example, uses this research design to study why states that undergo a revolution often end up soon thereafter in a war with their neighbors. By studying the foreign policies of states both before and after they undergo revolutions, Walt is able to use what is essentially a most-similar cases design to examine how a revolution changes the foreign policy of a state. Within the same study, Walt is also able to compare across revolutionary states to assess why some ended up in interstate wars and others did not. This allows Walt to get many different comparisons from studying three cases in detail (the French, Russian, and Iranian revolutions) and four cases in more abbreviated fashion (the American, Mexican, Turkish, and Chinese revolutions). It also permits Walt within a relatively small number of cases to have variance in both his independent variables (he argues that the difference between mass and elite revolutions is significant, with the former more likely to end in interstate war, and that defense dominance reduces the likelihood of revolutionary states engaging in war) and his dependent variable (some of his cases ended in interstate wars, and others did not).

Deviant Cases

Deviant cases are cases that do not conform to the predictions made by the theory or theories under investigation (George & Bennett, 2005, p. 240, note 14; Levy, 2002, p. 132). Put in Bayesian terms, deviant cases are those cases whose outcomes do not fit our prior theoretical expectations. Such cases can be identified as the cases in a statistical distribution that have high error terms vis-à-vis the model being tested or as cases in a small-*N* study

that do not fit the expectations of extant theories or that have outcomes different from those of similar cases. In principle, cases identified as deviant in either setting could defy expectations for a number of reasons: They may be the result of measurement error, of the combined effects of many left-out variables that individually have small effects on outcomes, of one or a few left-out variables that strongly influence outcomes, or of irreducibly stochastic or probabilistic processes. In practice, statistical researchers are often inclined toward a default assumption that outliers represent the effects of many weak variables or of stochastic processes, whereas case study researchers tend to assume that deviant cases result from measurement errors or small numbers of left-out variables that closer study might reveal (George & Bennett, 2005, pp. 20, 32).⁶ In this latter view, deviant cases can be analyzed at different times in the research cycle and are potentially powerful sources of new hypotheses and variables. Once identified through close analysis of a deviant case or cases, such new variables can be inserted back into large-*N* studies as a further check on their applicability.

Democratic peace theory is one well-developed research program in which substantial attention has been paid to deviant cases. Ray (1995, pp. 86-87) provides a list of 20 potential exceptions to the dyadic variant of the theory, cases where scholars suggest democracies have fought one another. Several of these have been studied as deviant case studies. For example, Miriam Fendius Elman (1997) investigates Finland's war with Great Britain during World War II. As Elman notes, "Finland was a democracy and retained its sovereignty during the war, but allied itself with an Axis power" (p. 191). Elman's analysis of the case provides an explanation into the reasons for its deviance and suggests fruitful lines of investigation to pursue for the wider research program. Democracies can be distinguished by their different political structures and, in particular, by the degree of autonomy that separate branches of government have over foreign policy choices. Finland had a semipresidential system, and the president had almost total power over what the country did in relation to its neighbors.⁷ Because the legislature had no control over foreign policy, the president was not required to "enlist widespread support before engaging in risky foreign policies that promise large-scale violence" (M. F. Elman, 1997, p. 197). More broadly, the results of the case study suggest that democratic peace theory may work better as a typological theory, with contingent generalizations based on different domestic political structures and the distribution of war/peace preferences across the different branches of government. Elman notes, for example, that, *ceteris paribus*, the theory's predictions are more likely to hold for decentralized

(e.g., coalitional parliamentary or presidential) democracies than centralized (e.g., semipresidential or Westminster) democracies.⁸

Although the study of deviant cases is potentially helpful, it may also lead to unfruitful theoretical amendments. If scholars respond to anomalies by simply relabeling them rather than providing an explanation for them, the new version of the theory will not be a real improvement (Lakatos, 1970). IR scholars have discussed at length how to judge iterations between theory and evidence.⁹ In particular, they are now more sensitive to what Alan Musgrave (1974, pp. 3, 7) calls the historical approach to confirmation of a scientific theory. The historical approach suggests that we cannot determine whether evidence supports a theory solely on the basis of whether it fits the current iteration of the theory. It is not enough to ask whether the theory covers known anomalies. It is also necessary to track the trajectory of a theory as it develops and ask whether amendments did more than just relabel empirical puzzles (C. Elman & Elman, 2002, 2003). The question is whether the new theory provided additional value, signaled by the prediction of novel facts.¹⁰ This is a key issue for qualitative methodologists, who often recommend that scholars move freely between theory and evidence, iteratively responding to puzzles the current variant of their theory cannot address (e.g., George & Bennett, 2005).

Conceptual Innovation

Qualitative approaches, especially the intensive study of one or a few cases, allow for the development of differentiated and more closely focused concepts. David Collier (1999) pointed to the special contributions of “researchers who are experts at ‘extracting new ideas at close range.’ These scholars are deeply engaged both with theory and with the close analysis of cases, giving them an unusual capacity to see the general in the particular” (p. 4). Although Collier was addressing country studies, his observation holds just as true for detailed case study knowledge arrived at through a full range of data collection methods, including archival research.

One example of conceptual development following from detailed knowledge of cases is how IR scholars have spent considerable effort debating the ways in which states generate military power and then how changes in relative power influence foreign policy choices and international political outcomes. Qualitative IR scholars, mainly political realists and their critics, have addressed various parts of this research. Collectively, these

scholars have developed more complete theoretical accounts covering and connecting the ways in which states produce and react to power.

For example, Stephen Peter Rosen's (1996) *Societies and Military Power: India and Its Armies* suggests that a state's military power is bounded by its social structures and by the connections between the armed forces and their host society. Differences in social structure, "the subunits of society with which people identify and to which they give their loyalties" (Rosen, 1995, p. 23), in turn will affect the amount of military power that can be generated from a given set of factor endowments. Rosen (1995) argues that "the dominant social structures of a group of people might lead to characteristic strengths and vulnerabilities of each society when making money or making war" (p. 24). States that have superior factor endowments can nevertheless be defeated if their divisive social structures prevent the effective generation and application of military capabilities.

Another example is Fareed Zakaria's (1998) *From Wealth to Power: The Unusual Origins of America's World Role*. Zakaria attributes the rise of America as an expansionist nation-state to the development of the mechanisms that allowed the government to extract and apply resources for military action. Between 1865 and 1890, except for the acquisition of Alaska and the annexation of the Midway Islands, the United States declined to apply its increasing societal wealth to expanding its territory. By contrast, between 1896 and 1917, America was highly expansionist, especially in Central America and the Caribbean. Zakaria (p. 8) attributes this change in state behavior to a change in the structure, scope, and capacity of the state. Although America's enormous material resources were previously wielded by a "decentralized, diffuse and divided" state the 1880s and 1890s saw a shift to a modern state better able to capture and direct those resources for a muscular foreign policy (pp. 11-12, 39-41).¹¹

With respect to changes in relative power, whether in terms of bids for hegemony and/or reactions to the rise of other states, the IR literature features a rich array of books utilizing case study methods that develop more sophisticated variants of concepts and their associated theories. For example, Victoria Tin-bor Hui's (2005) *War and State Formation in Ancient China and Early Modern Europe* contrasts China's Warring States period (656-221 B.C.E.) with Europe's Early Modern period (1495-1815 C.E.). Hui addresses the puzzle of why China's local balance of power systems failed but Europe's did not. Using a historical institutionalist framework, Hui eschews the search for universal generalizations, instead focusing on "causal mechanisms, which have varying effects, depending on contexts"

(p. 8). Hui follows McAdam, Tarrow, and Tilly (2001) in using “paired comparisons of uncommon cases to find out how recurrent causal mechanisms combine differently with varying initial and environmental conditions to produce radically different outcomes” (p. 8). By juxtaposing the logics of balancing and domination, in effect integrating both positive and negative feedback into her model, Hui is able to account for how the cases diverge.

Another recent example is Randall Schweller’s (2006) *Unanswered Threats: Political Constraints on the Balance of Power*, which investigates why states often do not balance in response to dangerous and unappeasable aggressors, when effective balancing is needed to deter or defeat those threats. Providing close and detailed case analyses that include Argentina and Brazil in the War of the Triple Alliance of 1864 to 1870, Britain and France between the world wars, and France from 1878 to 1898, Schweller argues that the more fragmented and diverse the state’s various elite and societal groups, the less it can be expected to respond appropriately to external strategic pressures. Schweller formalizes this insight into a four-variable model comprising elite consensus, elite cohesion, social cohesion, and regime vulnerability. The normal balance of power model suggests that elites will reach consensus on what responses are appropriate and then act to overcome internal mobilization hurdles.

Schweller contrasts this model with four underbalancing alternatives, each of which maps onto one of the extended case studies, where domestically generated pathologies interfere with efficient balancing. For example, in the additive model, a combination of social fragmentation, regime vulnerability, and elite fragmentation prevents the development of an elite balancing consensus, with the result that the state underbalances. The other three variants use the same basic building blocks but produce different causal narratives depending on complex conjunctions and temporal sequencing. In all four underbalancing models, the interaction between societal and elite problems prevents the nation-state from applying military capability as a rational response to external strategic constraints.

In each of these studies, the authors combine their theoretical concerns with a detailed knowledge of cases. Opening the black box and accounting for how states extract, apply, and react to changes in power allows for a more multifaceted appreciation of foreign policy and international politics. We suggest that this theory development would not have been possible without the authors’ “close, creative engagement with cases” (D. Collier, 1999, p. 5).

Typological Theories, Explanatory Typologies, and QCA/Fuzzy-Set Analysis

IR scholars have a long tradition of using explanatory typologies, including typologies of international systems and their associated stability or instability, types of states and their characteristic foreign policies, and types of individual leaders' personalities and their decision-making styles. Although scholars have traditionally been somewhat unselfconscious in this usage, there have been moves toward a more systematic approach.¹² Explanatory typologies can be distinguished from simple descriptive and classificatory typologies. Descriptive typologies build type-concepts from compounds of attributes. Here, the dimensions of the property space—the listing of all possible combinations of the attributes of interest—are inductively derived, as in early classifications of plants or animals. Each unique combination of attributes provides a separate compound concept. Explanatory typologies follow a similar conjunctural or combinatorial approach, except that the constituent attributes are extracted from the variables of a theory. The dimensions of the property space (its rows and columns) reflect alternative values of the theory's independent variables, so each cell in the space is associated with predicted values of the theory's intervening or dependent variables. Instead of asking "What constitutes this type?" explanatory typologies pose the question: "If my theory is correct, what do I expect to see for each combination of variables?"

Classificatory typologies determine to which type a case can be characterized as belonging. Beginning with a typology, empirical data are coded as falling into one cell or another, guiding scholars to answer the question "What is this a case of?" In explanatory typologies, the classificatory function exclusively focuses on evidence that can arbitrate the theoretical claims being made. For example, analysts may investigate a case to determine whether there is the anticipated congruence between its scores on the typology's dimensions and the predictions made for the cell in which the case is expected to belong. In addition, the analyst can use the location of cases in different cells as a guide to making the most productive comparisons for testing the underlying theory.

Schweller's (1998) *Deadly Imbalances* is an example of the explicit use of explanatory typologies in the IR subfield.¹³ Building on the classical realist tradition, Schweller's balance-of-interest theory is developed using an explanatory typology based on two dimensions: (a) the state's relative capabilities and (b) its interests, expressed as its view of the current distribution of power (i.e., whether it supports the status quo). Relative capabilities are

configured as the ordinal categories of poles, lesser great powers, midpowers, and small states, and the interests dimension is ordinally arranged into categories as well: unlimited-aims revisionists, limited-aims revisionists, indifference, support status quo/accepts limited revision, and strongly supports status quo. Schweller (1998, pp. 84-89) then takes the 20 cells produced by his two dimensions and compresses them using a combination of logical and pragmatic reduction techniques. Taken together, these compressions reduce the number of cells from 20 to 8 (see his Figure 3.1, p. 85). Within these cells, the different values of the theory's two variables (together with its constants, which include the anarchic international system and unit rationality) produce different predictions of state behavior.

Charles Ragin (1987, 2000) has developed a suite of methods that, in their use of property space and configurational causation, bear a strong family resemblance to typological theorizing, albeit with some important differences.¹⁴ Ragin's, 1987 book, *The Comparative Method*, emphasizes the configurational nature of cases and the prevalence of complex interaction effects. Qualitative comparative analysis (QCA) employs Boolean algebraic logic to analyze "dichotomously measured variables (i.e., 'crisp sets') and a veristic understanding of causation that does not systematically incorporate probabilistic considerations" (see Mahoney, 2004a, p. 2). Ragin's (2000) subsequent *Fuzzy-Set Social Science* allows scholars to use degrees of membership in categories rather than dichotomous codings of "in" or "out." For example, rather than giving nominal or ordinal measures of a case, a case may be measured as "fully in" the set of democracies (a score of 1.0), "mostly in" (a score of .75), "more in than out" (above .5), "mostly out" (.25), "fully out" (0), or somewhere in between these scores (Ragin, 2000; see also Goertz, 2006, on "family resemblance" concepts and measures). As Ragin argues, such fuzzy-set scores can offer advantages over crisp, nominal scores. For many intents and purposes, for example, once a state is fully democratic, it matters little if it is extremely democratic, and fuzzy-set scores accordingly set aside irrelevant variation among extreme cases.

Ragin's QCA/fuzzy-set approach has sparked a considerable debate among qualitative and other methodologists and some interest in applying both the crisp and/or fuzzy-set versions in the study of IR.¹⁵ This is one area, however, where the IR subfield has not paid as much attention as the comparative politics subfield.¹⁶ Although it has not yet been employed by a large number of scholars working within the qualitative tradition in the IR subfield, we anticipate that QCA/fuzzy set is likely to become significantly more prominent over time.

Process Tracing in IR Research

With process tracing, causation is established through uncovering traces of a hypothesized causal mechanism within the confines of one or a few cases. Cases may provide a variety of evidence of the operation of causal mechanisms, none of which is directly comparable, some of which may be more important than other pieces, and all of which taken together may allow analysts to draw conclusions about the adequacy or inadequacy of an explanation (Bennett & Elman, 2006a). Process tracing can involve both inductive and deductive study of events and sequences within a case. Inductive examination may reveal potentially causal processes that the researcher had not theorized a priori. Deductively, theories can suggest which intervening events should have occurred within a case if the theory is an accurate explanation of the case. Depending on the theory under investigation, some of the hypothesized steps in the case may be tightly defined necessary conditions, and others may be defined more loosely as having several substitutable processes that could have taken place at a particular juncture (on the distinction between necessary conditions and substitutable conditions, see Goertz, 2006).

There are many excellent examples of process tracing in IR research, including, but certainly not limited to, books by Drezner (1999), Eden (2004), Homer-Dixon (1999), Khong (1992), Knopf (1998), Larson (1997), Moravcsik (1998), Owen (1997), Rock (1989, 2000), Sagan (1993), Shafer (1988), Snyder (1984, 1991), Walt (1996), and Weber (1991).¹⁷ Each of these works shares, to some degree, a set of common features that constitute best practices in process tracing. These include: explicit attention to and process tracing on alternative explanations and on the hypothesis or explanation of most interest to the researcher; sustained focus on the question of “what else must be true” of the process through which the outcome arose if a proposed hypothesis explains the outcome; and relentless empirical research on these hypothesized processes, using a wide variety of sources (often including archived documents, contemporary news accounts, secondary histories, biographies or memoirs, and interviews) with due attention to the potential motivated and informational biases of each source.

Scott Sagan’s (1993) book, *The Limits of Safety: Organizations, Accidents, and Nuclear Weapons*, illustrates these virtues. Sagan set out to test two competing theories on the origins of accidents in complex organizations, the “normal accidents theory,” which envisions accidents as a normal development in even the most safety-conscious complex organizations, and the “high reliability theory,” which puts greater faith in the ability of organizations to provide

adequate interlocking safeguards against accidents. Sagan chooses to test these alternative explanations in the context of the potential accidental or unauthorized use of nuclear weapons. This poses a tough test for the normal accidents theory because nuclear weapons safety is a top priority, and nuclear weapons are managed by highly trained and disciplined military forces with adequate resources. This context also provides the opportunity for process tracing that strongly discriminates between the two theories, Sagan argues, as the specific conditions that promote safety in the view of the high reliability theorists can actually allow for interrelated and cumulating mistakes in the view of normal accidents theory.

Sagan (1993) spells out the predictions for each theory on the ways in which nuclear weapons accidents and near misses could arise or be prevented. He then tests these against the case of nuclear weapons safety in the Cuban missile crisis. Although his initial research on the basis of the available documentary record suggested that there were no close calls during the crisis, Sagan decided that available records were a potentially biased source because the same organizations responsible for nuclear weapons safety were also writing the reports on weapons safety during the crisis. Thus, Sagan interviewed individuals responsible for weapons safety during the crisis, including working-level personnel and higher-level decision makers, and he found that these individuals, mostly retired by the time he interviewed them, readily recounted several close calls during the crisis. These interviews helped Sagan conduct a more thorough review of the documentary evidence, including documents he obtained through Freedom of Information Act requests, and he was able to construct detailed histories of the close calls during the crisis from exhaustive research into these sources and memoirs, military archives, news accounts, congressional documents, and many other sources. Bolstering his analysis of the Cuba case with a briefer examination of other close calls in nuclear weapons safety, Sagan concludes that, on the whole, the normal accidents school better explains the near misses in nuclear weapons safety and hence provides a cogent warning for designing nuclear weapons safety procedures in ways that anticipate the “normal” dynamics through which accidents can arise.

Lynn Eden’s (2004) *Whole World on Fire: Organizations, Knowledge and Nuclear Weapons Devastation* investigates the consistent underestimation by the United States of the likely damage that would be caused by nuclear weapons. Its estimates of those consequences focused almost exclusively on blast effects, with very little attention paid to fire damage, which was in fact probably more destructive. This constant omission resulted in the United States procuring many more nuclear weapons than were needed to

achieve its strategic goals. Eden's explanation for the downplaying of fire damage is that the organization's knowledge was socially constructed. Powerfully influenced by the use of precision bombing during the Second World War, the United States paid much more attention to estimating blast than incendiary damage. As the Air Force developed more analytical tools to calculate blast, the gap between its knowledge and interest in blast and fire opened even further.

Eden's (2004) account fulfills several of the requirements for a convincing use of process tracing techniques. She uncovers substantial evidence that should have been found if her preferred explanation was to be held to be true. Using archives, interviews, and other qualitative data-collection procedures, Eden traces out the process by which the Air Force's knowledge of blast effects was first developed and then embedded in the organization and how fire damage was systematically ignored. In addition, Eden guards against confirmation bias by looking for evidence of a wide range of alternative explanations and finds evidence of observable implications that are inconsistent with them.¹⁸ Finally, Eden pays careful attention to the choice of where to begin and end her process trace.

A third example is Hendrik Spruyt's (2005) *Ending Empire: Contested Sovereignty and Territorial Partition*. Spruyt addresses the puzzle of why states have reacted so differently to the prospect of losing territories. A key element in his explanation for change in territorial policy is the number of veto points at the domestic level. Although Spruyt deductively develops a series of hypotheses from his institutionalist model, he avoids making overly universalist claims. As Spruyt observes "Any explanation of macro-level, complex historical events must be partial. . . . I lay no claim to a covering-law explanation that uniformly accounts for all features of territorial in all instances" (p. xii).

Spruyt (2005) tests his explanation by analyzing five primary cases: Britain and East Africa, France and Algeria, Holland and Indonesia, Portugal and its African possessions, and the dissolution of the Soviet Empire. Although this case-selection strategy allows for both diachronic and synchronic comparison, Spruyt is careful to note that rather than trying to develop nomothetic causal statements, he is instead explicating the causal mechanisms that can account for the observed events. "While *Ending Empire* takes its cue from deductive theorizing, the actual mechanisms through which actors operate, and the particular context and struggles of the political contest, must be traced by historical, inductive analysis" (p. 35, note 84).

Multimethod Works in IR

The IR literature includes many research designs that combine case study methods with formal models, statistical analysis, or both, including Biddle (2004), Fortna (2004), Goemans (2000), Huth (1988), Martin (1992), Schultz (2001), Reiter (1996), Kydd (2005), and Crawford (2003). Combining methods can be a particularly powerful approach, as each method can to some degree offset the limitations of other methods. In particular, researchers have used case studies to test whether the mechanisms derived from formal models or theorized to explain statistical distributions are in fact extant and observable in the processes through which outcomes have arisen in individual cases. Statistical analysis, in turn, can help identify outlier or deviant cases that do not fit a model very well, and these cases can be studied with the goal of identifying if measurement error or left-out variables help explain why they do not fit the model. Formal models can help make more rigorous the theories developed through the use of case studies and statistical analysis, leading to new and testable counterintuitive insights.

An excellent example of combining methods in this way is Lisa Martin's (1992) *Coercive Cooperation*. Martin uses a formal model to derive counterintuitive insights on economic sanctions, arguing that cooperation in imposing such sanctions should actually increase when the major state imposing the sanctions faces high costs and that the wider the number of states involved in imposing sanctions, the more that other states will want to join in enacting sanctions. Martin then not only statistically tests these arguments but uses individual case studies to show that the hypothesized mechanisms are indeed operative in individual cases.

Similarly, Hein Goemans (2000) uses a formal model to develop a hypothesis that semirepressive regimes will be strongly reluctant to lower their aims during a war, even if they learn from military engagements that they are likely to lose the war. This is because of the fact that the leaders of such regimes will pay a higher price than either democratic or authoritarian leaders if they lose a war. Goemans then statistically tests this hypothesis and also tests it using case studies of the decision making of several of the combatant countries in World War I. Like Martin, Goemans uses the case studies to determine if the hypothesized mechanisms are in fact operative in individual cases. This use of case studies to test or demonstrate the operation of hypothesized mechanisms is probably the most common purpose in combining case studies with other methods.

A second goal of combining methods is demonstrated in Dan Reiter's (1996) *Crucible of Beliefs: Learning, Alliances, and World Wars*. In this study of how the lessons leaders draw from their alliance experiences affect their subsequent alliance strategies, Reiter performed a statistical test of his hypothesis and then not only studied some individual cases that fit the hypothesis but also focused on the cases that did not fit the statistical model. From studying these individual cases, Reiter concludes that the presence or absence of democracy is an important variable omitted from his initial analysis, and he then performs additional statistical tests using this variable.

The benefits of juxtaposing statistical and case study methods are well illustrated in recent research on civil and ethnic conflicts, including the mechanisms behind individuals' participation in such conflicts and the structural conditions that make violence or reconciliation likely. As scholars began to focus renewed attention on civil and ethnic conflicts in the aftermath of the cold war, two prominent statistical studies argued that "greed," or opportunities for material or political gain through the instigation or perpetuation of conflict, was a more important factor in such conflicts than was "grievance," or clashes between identity groups over perceived unfair treatment (P. Collier & Hoeffler, 2001; Fearon & Laitin, 2003). Engaging these analyses, and drawing on a two-volume set of 21 case studies of civil wars he edited with Paul Collier (P. Collier & Sambanis, 2005a, 2005b), Nicholas Sambanis (2004) argues the need to get beyond a simple greed-grievance distinction. Instead, civil conflicts are, in his view, "the result of four interacting factors: the demand for loot, the demand for political change, the opportunity to mobilize criminal or insurgent groups, and the mechanisms (relational, emotional, cognitive, or environmental) that characterize claim making and resource extraction" (p. 269). Sambanis provides a cogent methodological critique of the earlier P. Collier and Hoeffler (2001) and Fearon and Laitin (2003) works, arguing that "by combining statistical and case study work we can better understand the political processes that lead societies to civil war" (Sambanis, 2004, p. 269). In much the same terms argued above, he maintains that case studies can help identify and clarify causal mechanisms, point out measurement problems in statistical studies, identify omitted variables, highlight interaction effects, indicate when unit heterogeneity exists, and recognize different mechanisms and types of civil war (Sambanis, 2004, pp. 259-260). Collectively, this and many other case studies of civil violence (e.g., Kalyvas, 2003; Posen, 1993; Ross, 2004) are contributing to a more complete if more complex understanding of civil violence that gets

beyond the simple distinction of greed *or* grievance to model how greed *and* grievance, in different combinations and interactions with lootable resources, geographic or international safe havens, and potential international interveners come together to produce qualitatively different paths to and kinds of civil violence.

Conclusion

As illustrated by these examples of deviant, least-likely and most-similar cases, conceptual innovation, typological theorizing, multimethod research, and process tracing, IR research has greatly benefited from the widespread use of qualitative methods. The recent resurgence in writings on qualitative methods suggests that this use is likely to become even more influential. Recent methodological texts have helped to clarify the best practices associated with qualitative approaches and to provide specific standards by which they should be judged. Perhaps even more important, however, has been the analysis of the epistemological foundations on which qualitative methods rely. The result has been a much clearer understanding of the comparative strengths and weaknesses of different tools and a more sophisticated appreciation of which standards are relevant to which methods. This promises an IR subfield in which qualitative methods are employed more rigorously and are both used and judged appropriately.

At the same time, there is ample room for continuing improvements in qualitative methods across the subfields of political science. Three areas, in particular, pose ongoing challenges and are likely to be the focus of new methodological innovations. First, although we noted above a few of the many good examples of multimethod research in IR, very little has been written about more general techniques and approaches for combining case studies with other methods.¹⁹ There are many possible combinations of case study, formal, statistical, survey, experimental, and other methods; and the characteristic strengths, weaknesses, and challenges of alternative combinations and sequences of multimethod work deserve examination. Second, technological changes are rapidly making it possible for qualitative researchers to provide other researchers with Internet access to their primary source materials, including documents, interviews, transcripts, audiotapes and videotapes, and photographs. Protocols for storing and providing access to these materials, and for making them searchable by appropriate terms and keywords, need to be developed. Such efforts can improve the replicability of qualitative research

and allow critics to assess researchers' interpretations of their qualitative sources. Finally, qualitative researchers need to continue to work on techniques for reliably assessing the identities, preferences, and perceptions of actors that are of interest to constructivists, rational choice theorists, and political psychologists.²⁰ Weeding out instrumental purposes of communication and getting actors to reveal their preferences and perceptions continue to pose difficult challenges, and in view of their long tradition of working closely with political actors, qualitative researchers are well positioned to make contributions toward these tasks.

Notes

1. For reviews of the role of case study methods in international political economy, international environmental issues, and international security issues, see, respectively, Odell (2004), Mitchell and Bernauer (2004), and Kacowicz (2004).

2. At great risk of leaving out important topics, these include the study of deterrence, coercive diplomacy, foreign policy decision making, crisis decision making, bureaucratic politics and organizational processes, the "democratic peace," alliance behavior, arms control, interactions between domestic and international politics, military intervention, economic sanctions, the causes of war, counterinsurgency warfare, environmental policy, military occupations, U.S.-Soviet relations, war termination, international institutions, military strategy, peacekeeping and conflict resolution, international trade, nuclear proliferation, transnational actors and movements, humanitarian intervention, terrorism, ethnic conflict and genocide, ideas, norms and foreign policy, and sovereignty. An earlier version of this note, subsequently edited for length, included more than 70 references of exemplar qualitative research from these subject areas.

3. On the comparative advantages and limitations of case study methods, see George and Bennett (2005, pp. 17-34); for more critical views of case studies, see Goldthorpe (1997), Lieberston (1992), and Njolstad (1990).

4. Although the selected works were almost all published within the past 10 years, some of them can also be viewed as late second-generation works in terms of the schema introduced in the introduction to this symposium.

5. Of course, a researcher might overlook differences between the cases that unknown rival hypotheses would emphasize, so like all methods of inference from observational data, this method is potentially fallible.

6. It should also be noted that a deviant case that is a "most likely" case (i.e., one that was strongly expected to conform to expectations generated by the theory but did not) is a valuable source of inferential leverage, albeit one that would cause us to question the theory's validity.

7. See M. F. Elman (1997, p. 196, note 11) for an explanation of semipresidential systems.

8. For another example of the analysis of a deviant case, see Joseph Grieco's (1995, 1996) investigation of the effect of the Maastricht Treaty on the European Union, which he argues is a deviant case for the neorealist research program. Grieco offered his "voice opportunity hypothesis" as a friendly amendment to realist theory that would allow it to explain the deviant case. He suggests that the new hypothesis allows neorealism to account for hitherto puzzling facts and point to new areas for promising research (Grieco, 1995, p. 40). Keohane and Martin

(2003, p. 95) suggest that Grieco's argument is actually institutionalist, not neorealist. Legro and Moravcsik (1999) make a similar criticism.

9. See, for example, the debates between John Vasquez and his critics, reproduced and extended in Vasquez and Elman (2003), and between Jeffrey Legro and Andrew Moravcsik (1999) and their respondents.

10. We should note that there is substantial disagreement on which definition of novelty to apply. See C. Elman and Elman (2002, 2003) for discussions of some of the potential standards, including strict temporal novelty, new interpretation novelty, heuristic novelty, and background theory novelty.

11. Two other good examples of studies investigating different aspects of the generation of state power are Christensen (1996) and Liberman (1996). Christensen (1996) argues that leaders can choose what appears to be an overly aggressive and confrontational grand strategy for the short term, incurring what appear to be unnecessary costs, if their underlying goal is to mobilize domestic support for long-term grand strategies. Liberman (1996) investigates whether power is a cumulating resource or whether conquered territory poses a net cost to invaders. He concludes that the larger surplus generated by a modern society, the more efficient coercive means available, and the fact that potential resisters have much more to lose all mean that conquered industrial societies may be profitable for their occupiers.

12. This section draws on C. Elman (2005).

13. See C. Elman (2004) for another example of an explicit explanatory typology within the family of realist research traditions; and see George and Bennett (2005) for discussion of Bennett, Leppgold, and Unger's (1997) analysis of burden sharing in the 1991 Persian Gulf War as an example of typological theorizing.

14. See Bennett and Elman (2006b) for a discussion of these differences.

15. For critiques and discussions see, for example, Lieberman (2004), Seawright (2004, 2005), Mahoney (2004b), Goertz and Mahoney (2005), and Bennett and Elman (2006b).

16. There have, however, been some applications in the international relations subfield. See, for example, Drezner (1999).

17. For brief descriptions of the research designs by Drezner (1999), Homer-Dixon (1999), Khong (1992), Knopf (1998), Larson (1997), Owen (1997), Shafer (1988), Snyder (1984), and Weber (1991), see George and Bennett (2005, pp. 194-197, 302-325).

18. As we note elsewhere (Bennett & Elman, 2006b, p. 460), for process tracing, eliminating competing explanations can be as important as finding positive evidence of the hypothesized causal mechanism. It might even be possible to have confidence in the existence of the hypothesized mechanism even without plentiful evidence of its operation, if key steps in all the alternative feasible explanations are shown to be untrue.

19. For a notable exception, see Lieberman (2005).

20. See Hopf (2002) for an outstanding example of qualitative research within the constructivist tradition.

References

- Bennett, A., & Elman, C. (2006a). Complex causal relations and case study methods: The example of path dependence. *Political Analysis*, 14, 250-267.
- Bennett, A., & Elman, C. (2006b). Qualitative research: Recent developments in case study methods. *Annual Review of Political Science*, 9, 455-476.

- Bennett, A., Leppgold, J., & Unger, D. (Eds.). (1997). *Friends in need: Burden-sharing in the Gulf War*. Basingstoke, UK: Macmillan.
- Biddle, S. (2004). *Military power: Explaining victory and defeat in modern battle*. Princeton, NJ: Princeton University Press.
- Christensen, T. J. (1996). *Useful adversaries: Grand strategy, domestic mobilization, and Sino-American conflict, 1947-1958*. Princeton, NJ: Princeton University Press.
- Collier, D. (1999). Letter from the president: Data, field work, and extracting new ideas at close range. *APSA-CP—Newsletter of the Organized Section in Comparative Politics of the American Political Science Association*, 10(1), 1-6.
- Collier, D., Mahoney, J. (1996). Insights and pitfalls: Selection bias in qualitative research. *World Politics*, 49(1), 56-91.
- Collier, P., & Hoeffler, A. (2001). Greed and grievance in civil war (2355). Washington, DC: World Bank.
- Collier, P., & Sambanis, N. (Eds.). (2005a). *Understanding civil war: Evidence and analysis, Vol. 1—Africa*. Washington, DC: World Bank.
- Collier, P., & Sambanis, N. (Eds.). (2005b). *Understanding civil war: Evidence and analysis, Vol. 2—Europe, Central Asia, and other regions*. Washington, DC: World Bank.
- Crawford, T. W. (2003). *Pivotal deterrence: Third-party statecraft and the pursuit of peace*. Ithaca, NY: Cornell University Press.
- Drezner, D. W. (1999). *The sanctions paradox: Economic statecraft and international relations*. New York: Cambridge University Press.
- Eden, L. (2004). *Whole world on fire: Organizations, knowledge, and nuclear weapons devastation*. Ithaca, NY: Cornell University Press.
- Elman, C. (2004). Extending offensive realism: The Louisiana Purchase and America's rise to regional hegemony. *American Political Science Review*, 98, 563-576.
- Elman, C. (2005). Explanatory typologies in qualitative studies of international politics. *International Organization*, 59, 293-326.
- Elman, C., & Elman, M. F. (2002). How not to be Lakatos intolerant: Appraising progress in IR research. *International Studies Quarterly*, 46, 231-262.
- Elman, C., & Elman, M. F. (2003). *Progress in international relations theory: Appraising the field*. Cambridge, MA: MIT Press.
- Elman, M. F. (Ed.). (1997). *Paths to peace: Is democracy the answer?* Cambridge, MA: MIT Press.
- Ember, C. L., Ember, M., & Russett, B. (1992). Peace between participatory politics. *World Politics*, 44, 573-599.
- Evangelista, M. (1999). *Unarmed forces: The transnational movement to end the cold war*. Ithaca, NY: Cornell University Press.
- Fearon, J., & Laitin, D. (2003). Ethnicity, insurgency, and civil war. *American Political Science Review*, 97, 75-90.
- Fortna, V. P. (2004). *Peace time: Cease-fire agreements and the durability of peace*. Princeton, NJ: Princeton University Press.
- Geddes, B. (2003). How the cases you choose affect the answers you get: Selection bias and related issues. In *Paradigms and sandcastles: Theory building and research design in comparative politics* (pp. 89-129). Ann Arbor: University of Michigan Press.
- George, A. L., & Bennett, A. (2005). *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.
- Goemans, H. E. (2000). *War and punishment: The causes of war termination and the First World War*. Princeton, NJ: Princeton University Press.

- Goertz, G. (2006). *Social science concepts: A user's guide*. Princeton, NJ: Princeton University Press.
- Goertz, G., & Mahoney, J. (2005). Two-level theories and fuzzy-set analysis. *Sociological Methods and Research*, 33, 497-538.
- Goldthorpe, J. H. (1997). Current issues in comparative macrosociology: A debate on methodological issues. *Comparative Social Research*, 16, 1-26.
- Grieco, J. (1995). The Maastricht Treaty, economic and monetary union, and the neorealist research programme. *Review of International Studies*, 21, 21-40.
- Grieco, J. (1996). State interests and international rule trajectories: A neorealist interpretation of the Maastricht Treaty and European economic and monetary union. *Security Studies*, 5, 176-222.
- Homer-Dixon, T. F. (1999). *Environment, scarcity, and violence*. Princeton, NJ: Princeton University Press.
- Hopf, T. (2002). *Social construction of international politics: Identities and foreign policies, Moscow 1955 and 1999*. Ithaca, NY: Cornell University Press.
- Hui, V. T.-B. (2005). *War and state formation in ancient China and early modern Europe*. New York: Cambridge University Press.
- Huth, P. K. (1988). *Extended deterrence and the prevention of war*. New Haven, CT: Yale University Press.
- Kacowicz, A. M. (2004). Case study methods in international security studies. In D. F. Sprinz & Y. Wolinsky-Nahmias (Eds.), *Models, numbers and cases: Methods for studying international relations* (pp. 107-125). Ann Arbor: University of Michigan Press.
- Kalyvas, S. (2003). The ontology of "political violence": Action and identity in civil wars. *Perspectives on Politics*, 1, 475-494.
- Katzenstein, P. J. (Ed.). (1996). *The culture of national security: Norms and identity in world politics*. New York: Columbia University Press.
- Keohane, R. O., & Martin, L. L. (2003). Institutional theory as a research program. In C. Elman & M. F. Elman (Eds.), *Progress in international relations theory: Appraising the field* (pp. 71-107). Cambridge, MA: MIT Press.
- Khong, Y. F. (1992). *Analogies at war: Korea, Munich, Dien Bien Phu, and the Vietnam decisions of 1965*. Princeton, NJ: Princeton University Press.
- King, G., Keohane, R. O., & Verba, S. (1994). *Designing social inquiry: Scientific inference in qualitative research*. Princeton, NJ: Princeton University Press.
- Knopf, J. (1998). *Domestic society and international cooperation: The impact of protest on U.S. arms control policy*. New York: Cambridge University Press.
- Kydd, A. H. (2005). *Trust and mistrust in international relations*. Ithaca, NY: Cornell University Press.
- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91-196). New York: Cambridge University Press.
- Larson, D. W. (1997). *Anatomy of mistrust: U.S.-Soviet relations during the cold war*. Ithaca, NY: Cornell University Press.
- Legro, J. W., & Moravcsik, A. (1999). Is anybody still a realist? *International Security*, 24(2), 5-55.
- Levy, J. S. (2002). Qualitative methods in international relations. In F. P. Harvey & M. Brecher (Eds.), *Evaluating methodology in international studies* (pp. 131-160). Ann Arbor: University of Michigan Press.
- Liberman, P. (1996). *Does conquest pay? The exploitation of occupied industrial societies*. Princeton, NJ: Princeton University Press.

- Lieberman, E. (2005). Nested analysis as a mixed-method strategy for comparative research. *American Political Science Review*, 99, 435-452.
- Lieberson, S. (1992). Small Ns and big conclusions: An examination of the reasoning in comparative studies based on a small number of cases. In C. Ragin & H. Becker (Eds.), *What is a case?* (pp. 105-118). New York: Cambridge University Press.
- Lieberson, S. (2004). Comments on the use and utility of QCA. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2(2), 13-14.
- Mahoney, J. (2004a). Introduction: A note on terminology. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2(2), 2.
- Mahoney, J. (2004b). Reflections on fuzzy-set/QCA. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2(2), 17-21.
- Mahoney, J., & Goertz, G. (2006). A tale of two cultures: Contrasting quantitative and qualitative research. *Political Analysis*, 14, 227-249.
- Maoz, Z. (2002). Case study methodology in international studies: From storytelling to hypothesis testing. In F. P. Harvey & M. Brecher (Eds.), *Evaluating methodology in international studies* (pp. 161-186). Ann Arbor: University of Michigan Press.
- Martin, L. L. (1992). *Coercive cooperation: Explaining multilateral economic sanctions*. Princeton, NJ: Princeton University Press.
- McAdam, D., Tarrow, S., & Tilly, C. (2001). *Dynamics of contention*. New York: Cambridge University Press.
- Mitchell, R., & Bernauer, T. (2004). Beyond story-telling: Designing case study research in international environmental policy. In D. F. Sprinz & Y. Wolinsky-Nahmias (Eds.), *Models, numbers and cases: Methods for studying international relations* (pp. 81-106). Ann Arbor: University of Michigan Press.
- Moravcsik, A. (1998). *The choice for Europe: Social purpose and state power from Messina to Maastricht*. Ithaca, NY: Cornell University Press.
- Musgrave, A. (1974). Logical versus historical theories of confirmation. *British Journal for the Philosophy of Science*, 25, 1-23.
- Njolstad, O. (1990). Learning from history? Case studies and the limits to theory-building. In O. Njolstad (Ed.), *Arms races: Technological and political dynamics* (pp. 220-246). Newbury Park, CA: Sage.
- Odel, J. S. (2004). Case study methods in international political economy. In D. F. Sprinz & Y. Wolinsky-Nahmias (Eds.), *Models, numbers and cases: Methods for studying international relations* (pp. 56-80). Ann Arbor: University of Michigan Press.
- Owen, J. M. (1997). *Liberal peace, liberal war: American politics and international security*. Ithaca, NY: Cornell University Press.
- Posen, B. (1993). The security dilemma and ethnic conflict. *Survival*, 35, 27-47.
- Ragin, C. C. (1987). *The comparative method: Moving beyond qualitative and quantitative strategies*. Berkeley: University of California Press.
- Ragin, C. (2000). *Fuzzy-set social science*. Chicago: University of Chicago Press.
- Ray, J. L. (1995). *Democracy and international conflict: An evaluation of the democratic peace proposition*. Columbia: University of South Carolina Press.
- Reiter, D. (1996). *Crucible of beliefs: Learning, alliances, and world wars*. Ithaca, NY: Cornell University Press.
- Rock, S. R. (1989). *Why peace breaks out: Great power rapprochement in historical perspective*. Chapel Hill: University of North Carolina Press.

- Rock, S. R. (2000). *Appeasement in international politics*. Lexington: University Press of Kentucky.
- Rosen, S. P. (1995). Military effectiveness: Why society matters. *International Security*, 19(4), 5-31.
- Rosen, S. P. (1996). *Societies and military power: India and its armies*. Ithaca, NY: Cornell University Press.
- Ross, M. (2004). How do natural resources influence civil war? Evidence from thirteen cases. *International Organization*, 58(1), 35-67.
- Sagan, S. D. (1993). *The limits of safety: Organizations, accidents, and nuclear weapons*. Princeton, NJ: Princeton University Press.
- Sambanis, N. (2004). Using case studies to expand economic models of civil war. *Perspectives on Politics*, 2, 259-279.
- Schultz, K. A. (2001). *Democracy and coercive diplomacy*. New York: Cambridge University Press.
- Schweller, R. L. (1998). *Deadly imbalances: Tripolarity and Hitler's strategy of world conquest*. New York: Columbia University Press.
- Schweller, R. L. (2006). *Unanswered threats: Political constraints on the balance of power*. Princeton, NJ: Princeton University Press.
- Seawright, J. (2004). Qualitative comparative analysis vis-à-vis regression. *Qualitative Methods: Newsletter of the American Political Science Association Organized Section on Qualitative Methods*, 2(2), 14-17.
- Seawright, J. (2005). Qualitative comparative analysis vis-à-vis regression. *Studies in Comparative International Development*, 40(1), 3-26.
- Shafer, M. (1988). *Deadly paradigms: The failure of U.S. counterinsurgency policy, 1945-1965*. Princeton, NJ: Princeton University Press.
- Snyder, J. L. (1984). *The ideology of the offensive: Military decision making and the disasters of 1914*. Ithaca, NY: Cornell University Press.
- Snyder, J. L. (1991). *Myths of empire: Domestic politics and international ambition*. Ithaca, NY: Cornell University Press.
- Spruyt, H. (2005). *Ending empire: Contested sovereignty and territorial partition*. Ithaca, NY: Cornell University Press.
- Vasquez, J. A., & Elman, C. (Eds.). (2003). *Realism and the balancing of power: A new debate*. Upper Saddle River, NJ: Prentice Hall.
- Walt, S. M. (1996). *Revolution and war*. Ithaca, NY: Cornell University Press.
- Weber, S. (1991). *Cooperation and discord in U.S.-Soviet arms control*. Princeton, NJ: Princeton University Press.
- Zakaria, F. (1998). *From wealth to power: The unusual origins of America's world role*. Ithaca, NY: Cornell University Press.

Andrew Bennett is a professor of government at Georgetown University. He is the coauthor, with Alexander George, of *Case Studies and Theory Development in the Social Sciences*, which won the Giovanni Sartori Prize for the best book on or using qualitative methods in 2005. He is also president and cofounder of the Consortium on Qualitative Research Methods. He is currently at work on a book on the politics and psychology of blame over failed predictions regarding the conflict in Iraq.

Colin Elman is an associate professor of political science at Arizona State University. He is currently engaged on a book for Cambridge University Press investigating the rise of the

United States to regional hegemony. He is one of the founders of both the International History and Politics and the Qualitative Methods organized sections of the American Political Science Association and a cofounder and executive director of the Consortium for Qualitative Research Methods.