International Studies Quarterly

International Studies Quarterly (2013) 1–13

Reducing Uncertainty: Information Analysis for Comparative Case Studies[1]

Katya Drozdova
Seattle Pacific University

and

Kurt Taylor Gaubatz
Old Dominion University

The increasing integration of qualitative and quantitative analysis has largely focused on the benefits of in-depth case studies for enhancing our understanding of statistical results. This article goes in the other direction to show how some very straightforward quantitative methods drawn from information theory can strengthen comparative case studies. Using several prominent "structured, focused comparison" studies, we apply the information-theoretic approach to further advance these studies' findings by providing systematic, comparable, and replicable measures of uncertainty and influence for the factors they identified. The proposed analytic tools are simple enough to be used by a wide range of scholars to enhance comparative case study findings and ensure the maximum leverage for discerning between alternative explanations as well as cumulating knowledge from multiple studies. Our approach especially serves qualitative policy-relevant case comparisons in international studies, which have typically avoided more complex or less applicable quantitative tools.

The epic methodological battles of the late twentieth century have largely subsided in light of the eminently reasonable notion that there are benefits to be gained from both the empirical confidence that comes from broad aggregate studies and the in-depth understanding generated by more focused case studies (Coppedge 1999). This reconciliation has brought a rising interest in the use of "multi-methods" to pair quantitative and qualitative work in the analysis of particular problems (Lieberman 2005). The multi-methods approach has primarily focused on the parallel application of large-N and small-n analytics to the same empirical issue. In this paper, we argue for an even tighter integration of quantitative and qualitative methods and demonstrate a quantitative but simple and accessible approach to enhance small-n case study research.

Our proposed approach applies where traditional statistics fall short. It complements and offers unique advantages over existing quantitative tools for small-n studies. Most importantly, we aim to aid qualitative scholars who typically would not use quantitative tools, but would benefit significantly from these improvements. To this end, we draw on information theory to propose a rigorous yet simple and broadly accessible approach to uncertainty reduction (Shannon 1948; Cover and Thomas 2006). We especially focus on policy-relevant comparative case studies involving assessments of the relative impacts of multiple factors theorized to affect an uncertain outcome—under the constraints of few cases, significant challenges in gathering comparable data, and potentially very consequential policy implications of decisions that may be informed by such studies. Our approach is thus in the tradition of, and aims to strengthen, the case study methods frequently used in international studies and explicitly designed for generating and accumulating policy-relevant knowledge across multiple cases.

The "structured, focused comparison" is a leading such method articulated by Alexander L. George (1979) as a response to the heightened interest in more systematic applications of qualitative methods to derive policy-relevant empirical generalizations (Eckstein 1975). In structured, focused comparison, a set of theoretically motivated critical variables is identified and then their variation is analyzed across several detailed qualitative case studies to derive systematic conclusions. While the structured, focused comparison method is admirably systematic and analytic, we argue that its empirical applications have often failed to provide analytic clarity. Drawing on recent advances in the field of information theory, we propose a straightforward method to provide a systematic quantitative understanding of the strengths and limitations of structured, focused comparisons and to reduce the uncertainty often associated with their results.

We draw on information theory, which tackles uncertainty, to offer analytic tools where traditional statistics typically fall short in the very small-n world of comparative case studies. The small number of observations limits the applicability and effectiveness of typical statistical tests such as correlation or regression analysis. The information-theoretic approach makes no assumptions about the underlying distribution of data and thus is not limited by the Normal distribution assumption of many traditional

[1] *Author's notes*: The authors would like to thank Dr. Michael Samoilov, of UC Berkeley's California Institute for Quantitative Biosciences (QB3), as well as the *ISQ* reviewers for helpful comments and suggestions. K. Drozdova would like to gratefully acknowledge the Hoover Institution at Stanford University for the opportunity to serve as a visiting scholar and benefit from research in the Hoover Institution Library and Archives while writing this article. The data, the worked out examples, an annotated Excel demonstration, and an R script file are available at kktg.net/kurt/entropy.

Drozdova, Katya and Kurt Taylor Gaubatz. (2013) Reducing Uncertainty: Information Analysis for Comparative Case Studies. *International Studies Quarterly*, doi: 10.1111/isqu.12101

statistics, nor situations where the distribution is unknown. The proposed approach is designed to help *qualitative* researchers systematically assess and reduce the uncertainty associated with their small-n samples of cases and the complexity of factor interactions explored by the in-depth case studies. The method uses the case study data to enhance the qualitative analysis results by offering ways to systematically rank order and compare the impacts of multiple factors that emerge from the qualitative analysis, with the help of some intuitive and simple calculations.

To present this approach, we begin with a brief review of the method of structured, focused comparison, and its strengths and limitations. Next, we outline the information-theoretic method, contrasting it with other tools to show how it advances comparative case study with complementary tools aimed to help qualitative scholars especially in international studies. We demonstrate the information-analytic process by applying it to three prominent examples of structured, focused comparison in international studies. In each example, the information theory approach sharpens the analysis by providing quantitative measures of uncertainty reduction and the relative impact of multiple theoretic factors identified by the qualitative studies on the outcomes they explored using structured, focused comparison. The findings enhance the understanding of comparative results and their policy implications as well as generate suggestions for advancing future case study research with information analytics.

The Strengths and Limits of Structured, Focused Comparison

The method of structured, focused comparison seeks to make case studies scientific. It integrates the advantages of qualitative methods with systematic procedures and the potential for generalizability typically associated with large-N statistical studies. George and Bennett explain the straightforward logic of the method:

> The method is "structured" in that the researcher writes general questions that reflect the research objective and that these questions are asked of each case under study to guide and standardize data collection, thereby making systematic comparison and cumulation of the findings of the cases possible. The method is "focused" in that it deals only with certain aspects of the historic cases examined (George and Bennett 2005:67).

In a typical structured, focused comparison, researchers identify the research problem and the variables of interest for that problem. They then use a systematic theory-driven process to select a set of relevant cases. These can be either cross-sectional or temporal in terms of different slices from a single case in which there is variation in the dependent variable at different points in time. Variation in the explanatory variables and the outcomes are then qualitatively assessed to identify the most important factors.

The method of structured, focused comparison has been endorsed by a wide spectrum of political scientists interested in qualitative methods. Van Evera (1997), for example, advocates the structured, focused approach, arguing that basic principles of the scientific method ought to apply to case studies in the social sciences. The same theme underlies King, Keohane, and Verba's *Designing Social Inquiry* (1994). Carlsnaes (1992) points to structured, focused comparison as a particularly appropriate tool for constructivist scholarship.

The goal of making qualitative case studies structured, focused, and more systematic can be further enhanced with the basic tools of information analysis. Structured, focused comparison is a sufficient analytic tool when a set of cases clearly aligns to distinguish the impact of one or two central variables. As demonstrated by even some of the most prominent structured, focused comparison studies, such clarity is rare. Just as in large-N regression methods, we need a rigorous and replicable way to assess the relative explanatory power of the different factors and their ability to clarify results, yet in a way that is accessible and useful for qualitative scholars. Information analytics enable greater clarity and metrics for assessing the impacts of multiple variables interacting in complex and uncertain ways, where a simple visual comparison of the results alone risks being insufficient or potentially misleading.

Three prominent examples of structured, focused comparisons illustrate these challenges and serve as a useful test bed for the meta-analytic approach we propose: *The Limits of Coercive Diplomacy* by Alexander George and William Simons (George, Hall, and Simons 1971; George, Simons, Gordon, Sagan, and Zimmerman 1994), *The Politics of Arms Control Treaty Ratification* by Michael Krepon and Dan Caldwell (1991), and *The Evolution of Peacekeeping* by William Durch (1993). George and Simons as well as Krepon and Caldwell, in particular, have been cited as exemplars of the structured, focused methodology (George and Bennett 2005). Surprisingly, all three of these important studies stop short of providing clear conclusions about the comparative impact and importance of the factors around which they are structured and focused. The Durch volume, most egregiously, runs through twenty cases, but provides no synthesis or overview of the results. In the other two studies, the authors provide an overview of the cases but do not provide clear guidance on the relative impact of the factors they are studying on case and policy outcomes. Each of the latter two studies provides a table summarizing the presence or absence of factors (independent variables) examined in the cases, but, oddly, does not include the case outcomes (dependent variable values).

Moreover, the actual analysis of the tables is largely left to the reader's intuition. The problem is conveyed in Table 1, which displays the data from *The Limits of Coercive Diplomacy*, with our addition of a "Success" row showing case outcomes. (In this and subsequent such data tables, the columns show cases and the rows show factors or variables analyzed for each case by the original studies.) The two unambiguously successful cases in Table 1, Laos and Cuba, have positive values for most of the independent variables. This does not help us understand the importance of the individual factors, many of which are also present in the unsuccessful or ambiguous cases. This example also suggests that the results might be highly dependent on certain cases or a subset of the variables, requiring further analytic tools to examine such potential dependence.

Krepon and Caldwell's study of arms control treaty ratification is similarly ambiguous. Their results are presented here as Table 2—with the addition of a "Ratification" row showing case outcomes, which, again, the original study strangely lacks. This missing representa-

Table 1. The Limits of Coercive Diplomacy

	Pearl Harbor	*Laos*	*Cuba*	*Vietnam*	*Libya*	*Nicaragua*	*Persian Gulf*
Success	N	Y	Y	N	A	Y	N
Clarity of objective	+	+	+		?		+
Strong motivation	+	+	+	+	+	+	+
Asymmetry of motivation		+	+		?		
Sense of urgency	+	+	+				?
Strong leadership		+	+	+	+	+	+
Domestic support	+	?	+		+		+
International support	+	+	+				
Fear of unacceptable escalation		+	+		?	?	+
Clarity of terms	?	+	+				

(*Notes*. Data from George and Simons, The Limits of Coercive Diplomacy (1994:288) with the addition of "Success" row showing case outcomes as described by the book's case studies. "N" refers to "No", "Y" to "Yes"; "A" to "Ambiguous", "+" indicates the presence of the row's factor in the corresponding column's case; "?" means that it is not clear whether the factor is present).

tion is further indicative of the difficulty of cumulative analysis in qualitative work. Krepon and Caldwell are forthright in describing the challenge of drawing conclusions from their seven case studies as "daunting" (1991:399). Their assessment of the structured, focused method is that it has *not* "provided clarity as to the rank ordering for the most important components of success or failure in the cases" (1991:400). Our proposed method rectifies this problem.

Durch's *The Evolution of UN Peacekeeping* (1993) is explicitly designed as a structured, focused comparison, but there is neither an overview of the variables of interest nor a summary of the findings across the cases. Reading through the cases makes a set of case questions reasonably clear. We have assembled a data set from our reading of the case descriptions. Our version of Durch's data is displayed in Table 3.

Despite being some of the most prominent examples of the method of structured, focused, comparison, none of these studies present their results in a way that allows for clear assessment of the relative impact of the posited factors on outcomes or for assessing the comparative findings. The two studies that do provide tabular summaries of their comparative results fail to include the case outcomes. The UN peacekeeping study lacks any such cross-case summary tabulation.

These exemplars reflect the difficulty of assessing the complex interactions between variables and outcomes, leaving much uncertainty about their mutual impact despite careful, in-depth, substantive, and qualitative analyses of each case. The human brain is very effective at grasping qualitative information, but systematic comparisons become increasingly difficult with more factors. The simple step of assembling both the explanatory and the outcome data in one table greatly benefits the visual and intuitive comparison of the results. But to really assess the patterns contained in these data, we need a more systematic and reproducible methodology. We propose here the use of information theory to clarify such patterns and thus leverage the structured, focused case study method.

Table 2. Arms Control Treaty Ratification

	Versailles Treaty	*Washington Naval Treaties*	*Geneva Protocol (1926)*	*Limited Test Ban Treaty*	*ABM*	*SALT II*	*INF*
Ratification	N	Y	N	Y	Y	N	Y
Perception of substantive treaty benefits	+	+	+	+	+		+
Presidential popularity	+	+		+	+		+
Perception of president as defender of U.S. national security interests		?		+	+		+
Perception of president as experienced in foreign affairs					+		
Presidential skill in handling executive-congressional relations		+		+	+		+
Quality of presidential advice	+	+		+	+	+	+
Favorable international environment	+	+	+	+	+		+
Support of senate leadership and pivotal senators		+		+	+		+
Support of military leadership	+	+		+	+	+	+

(*Notes*. Data from Krepon and Caldwell, *The Politics of Arms Control Treaty Ratification* (1991:465) with the addition of "Ratification" row showing case outcomes as described by the book's case studies. "N" refers to "No", "Y" to "Yes"; "+" indicates the presence of the row's factor in the corresponding column's case; "?" means that it is not clear whether the factor is present).

Table 3. Basic Data from *The Evolution of UN Peacekeeping*

	UNSCOB	UNTSO	UNEF I	UNEF II	UNDOF	UNOGIL	UNIFIL	UNYOM	UNFICYP	UNIIMOG	UNIKOM	UNMOGIP	UNTEA	UNGOMAP	ONUC	UNTAG	UNAVEM I	UNAVEM II	MINURSO	ONUCA
Success	Y	N	Y	Y	Y	N	N	N	Y	Y	Y	N	Y	Y	N	Y	Y	N	N	Y
Local Consent			+	+	+						+		+	+		+	+			+
Great Power Support	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+	+		+
United States Support	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+	+	+		+
Attitude Change			+	+	+						+		+	+		+	+			+
Domestic Conflict	+					+	+	+	+					+	+	+	+	+		+
International Conflict	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+		+	+
UNSC Initiative	+	+		+	+		+			+	+			+		+			+	
Local Initiative	+					+		+	+			+	+		+		+	+		+
Third-Party Broker			+		+			+	+				+							+
Political Mission	+					+	+		+				+		+	+		+	+	
Military Mission		+	+	+	+		+	+	+	+	+	+	+	+	+	+	+	+	+	+
Narrow Mandate			+	+	+	+		+			+	+	+	+		+	+	+	+	+
Broad Mandate	+	+					+		+	+					+					+
Mandate Revision		+		+					+			+			+	+		+		+

(*Notes.* "+" indicates the presence of the row's factor in the corresponding column's case; an empty space indicates otherwise).

This approach can provide systematic, comparable, and replicable measures of uncertainty and influence for the identified factors.

Information Theory and Political Science

Information theory emerged from the study of communications to answer fundamental questions about information transfer in uncertain (noisy) environments where people might lack the capacity to understand messages or where messages might become distorted in transmission. Communication experts in the 1940s thought that increasing the transmission rate of information over a communication channel increased the probability of error or distortion. However, Claude Shannon proved that this was not true as long as the communication rate was below channel capacity. The capacity was simply computable from the noise characteristics of the channel. Moreover, Shannon argued that random processes such as speech or music had an irreducible complexity (below which the signal could not be compressed any more), which he named information entropy (in parallel to entropy in thermodynamics), and argued that if the entropy of the source was less than the capacity of the channel, then asymptotically error-free communication was possible (Cover and Thomas 2006).

Following early applications to communications and cryptography (Shannon and Weaver 1949), information theory has made fundamental contributions to many fields, including physics, statistics, molecular biology, analytical chemistry, and finance (Verdú 1998; Guizzo 2003; Cover and Thomas 2006) as well as the social sciences. Recent advances have used information-theoretic methods to study complex living systems—including natural (Samoilov 1997; Samoilov, Arkin, and Ross 2001) and social systems (Drozdova 2008; Drozdova and Samoilov 2010).

The insights in Shannon's 1948 paper, which included the first use of the term "bit" to refer to a piece of data, paved the way for the revolution in digital communications (Guizzo 2003).[2] Looking back in 1990, Scientific American labeled the paper "the Magna Carta of the information age" (Horgan 1990). In addition to providing the foundation for nearly everything that followed in the development of digital communications, storage, and processing, Shannon's insights provided a new approach to assessing the information extant in other kinds of noisy data. A centerpiece of Shannon's contributions was the idea that analyzing noise could be informative about the signal.

Despite the increasing use of information analytics across this wide range of fields, these methods have seen relatively little use in political science. The basic intuition of information analysis is simply to provide a measure of how much knowing about the presence or absence of a given factor reduces uncertainty about the presence or absence of a given outcome. Since the data we deal with in political science and international studies are inevitably noisy due to the probabilistic and random elements that are inherent to human behavior, information theory should be able to make strong contributions in this domain as well. Political scientists have appropriately recognized Shannon's original work on information and

[2] Not coincidentally, perhaps, 1948 was also the year the invention of the transistor was unveiled by Bell Laboratories, where Shannon also worked. James Gleick (2011) argues that Shannon's paper was the more important of the two events.

communication—in the introduction to his 1984 Presidential Address to the American Political Science Association, Philip Converse called it a "watershed study" (1985). Those few political scientists who have drawn on information theory have used it as a way of understanding political information and communication processes (Lowry and Marr 1975; Oppenheim 1978; Congleton 2001). Others have used it creatively to understand uncertainty (Midlarsky 1974), issue diversity (McCombs and Zhu 1995), party structure (Dodd 1974; Molinar 1991), and other systemic characteristics (Rapoport 1974; Sheingate 2006).

Drozdova (2008) suggests the use of information theory for case study methods in the study of how organizations use technology and human networks to survive in different environments. In that study, based on a theory-driven sample of organizations and factors, the information analysis determined the two jointly most informative variables reflecting organizational missions and environments, which served as dimensions to select contrasting organization cases for in-depth comparative investigation of their alternative survival strategies. Here, we build on that approach to generalize the use of information theory as a meta-analytic for enhancing qualitative case study methods.

Uncertainty Analysis in International Studies: An Information-Theoretic Approach

As shown in the prominent examples of structured, focused comparisons discussed above, this well-established and otherwise qualitatively rigorous method loses rigor when the assessment of the results is left to the reader's intuition. It is difficult to grasp and compare many factors at a glance, leaving uncertainty about the relative importance of case factors and results and thereby creating opportunities to gain more knowledge from the information generated by case studies.

The information-theoretic approach assesses and reduces this uncertainty by *quantitatively* answering the questions:

1. What is the magnitude of uncertainty?
2. How much can we learn about an uncertain outcome given an observed factor that is theorized to be related to the outcome?
3. What are the relative impacts of multiple factors on the outcome?
4. What factors are most useful to know in order to reduce the outcome uncertainty?
5. What is the impact of the individual cases on the analytic conclusions?

Information theory provides the tools to answer such questions systematically using probabilistic measures of uncertainty and information that one variable contains about another.

This approach can be particularly valuable for policymakers, who must almost always make decisions under great uncertainty and with limited resources for exploring a situation. Case studies tend to point to a number of possibly important factors, but which one(s) should be prioritized and why? We need to know which of the many possible variable factors will give us the greatest information gain about the likely outcome. This knowledge can help inform and improve policy decisions.

Mutual information is a measure of how the knowledge of a given factor reduces the uncertainty about the outcome. It does not necessarily eliminate uncertainty, which may be impossible, but rather provides a systematic understanding of the information and uncertainty that exist within a structured case analysis. Higher mutual information indicates greater reduction in uncertainty—or greater knowledge gain. A variable with higher mutual information may be interpreted as having greater explanatory or predictive power about the outcome relative to other independent variables analyzed. This approach is different from and less restrictive in assumptions than, for example, using a regression analysis. We are not positing and testing a particular quantitatively modeled relationship, but rather estimating the knowledge content of the information generated by the qualitative case studies. This somewhat more modest goal is more applicable and accurate when studying phenomena marked by small numbers of observations and unknown factor distributions.

Frequencies of factor and outcome occurrences are used to estimate probabilities, so the necessary probabilities are easily calculated by simply counting the number of occurrences and co-occurrences of the independent and dependent variable values (demonstrated below). The analysis finds and systematically compares the entropy due to each factor and identifies each factor's relative impact in reducing entropy—that is, reducing the outcome uncertainty based on the observed factors (Samoilov 1997; Dhar, Chou, and Provost 2000; Provost and Fawcett 2001). Numerical results allow unambiguous ranking and systematic evaluation of each factor's relative impact on outcome uncertainty. Ultimately, this approach provides a powerful and stochastically rigorous way of understanding and improving case comparison results. At the same time, it is a method that is intuitively and mathematically accessible to a wide range of researchers. If you can count, you can do this.

The Uncertainty Reduction Method

The method we propose uses calculations based on simple frequency counts. It can be summarized in four steps —quantify, count, compute, and compare:

Step 1: Quantify

Assign 1 to positive variable values, 0 otherwise, to quantify case study data.

We demonstrate the use of this binary quantification for the three prominent examples of structured, focused comparison discussed above (Tables 1–3). For each study, the resulting set of mutually exclusive binary values for each set of cases then serves as data for the information analysis in Steps 2–4.

Step 2: Count

Count the number of times each of the outcome and factor values occur together to estimate conditional probabilities based on factor and outcome frequency counts.

A joint occurrence of a factor X and outcome Y is written (x, y). A binary system has four possibilities: $(x = 0, y = 0)$, $(x = 1, y = 0)$, $(x = 0, y = 1)$, $(x = 1, y = 1)$.

We count the number of times each of these possibilities occurs in a given case study and use these simple frequency counts for probability calculations. Conditional probability of outcome y given the value of factor x_i is calculated by dividing the joint probability that both values co-occur, written $p(x, y)$, by the probability that this factor occurs at all, $p(x_i)$, for all possible combinations (Table 7 in the methodological appendix shows an example of all calculations step by step):

1. Count factor and outcome co-occurrence, written $count(x, y)$, for each combination: $count\ (x_i = 0, y = 0)$, $count\ (x_i = 1, y = 0)$, $count\ (x_i = 0, y = 1)$, $count\ (x_i = 1, y = 1)$. Calculate joint probabilities, written $p(x, y)$, by dividing each count by the total number of cases, n. Thus, the calculation is $p(x, y) = count(x, y)/n$ for each factor and outcome combination.
2. Calculate probability, written $p(x)$, of each factor, where the probability of factor presence $p(x_i = 1)$ is found by dividing the count of positive occurrence over the total number of cases, written $count\ (x_i = 1)/n$. The complementary probability of factor absence, $p(x_i = 0)$, is $count(x_i = 0)/n$ or simply $1-p\ (x_i = 1)$ since total probability adds up to one.
3. Calculate conditional probabilities, written $p(y|x)$, by dividing each joint probability by each factor probability, $p(y|x) = p(x, y)/p(x)$.

Step 3: Compute

Plug the probabilities from Step 2 into the information analysis formulas.

Using notation where X means a factor (an independent variable) and Y means an outcome theorized to be related to that factor (a dependent variable):

1. Information entropy measures uncertainty of Y, written $H(Y)$.
2. Conditional entropy measures uncertainty of Y given X, written $H(Y|X)$.
3. Mutual information measures the reduced uncertainty in Y due to the knowledge of X, written $I(Y; X) = H(Y) - H(Y|X)$, meaning the difference between entropy (uncertainty about the outcome in the absence of any additional information) and conditional entropy (uncertainty about the outcome given that we know whether a factor is present or not) (Shannon 1948; Shannon and Weaver 1949; Cover and Thomas 2006).[3]

All of these probabilities are calculated from the simple counts collected in Step 2. These count based probabilities are then the basis for the entropy metric. Shannon's central insight was that the probability of joint events could be translated from a multiplicative to a more intuitive additive measure through the use of logarithms. Adding up the logarithmic measures would produce a simple concave function that shows maximum uncertainty where the probability of the outcome is 0.5 (a fifty-fifty chance) and minimum uncertainty where the probability of the outcome is either one or zero. The base of the logarithm corresponds to the units in which the uncertainty would be expressed. Hence, a base 2 logarithm would produce a binary measure of information and uncertainty. It was this approach that paved the way for the binary encoding of complex information.

[3] Conditional entropy and mutual information can be calculated in more complex ways based on several independent variables at a time as well as in terms of sequences or other combinations of variables. However, we introduce the method using simple binary calculations with our purpose in mind of providing simple but useful tools for qualitative scholars.

For our purposes in understanding the contributions of a given factor to knowledge about an outcome, the uncertainty measure varies from zero when the independent variable perfectly co-occurs with the dependent variable and as a concave function rises at a decreasing rate as joint occurrences decline.

The appendix provides the straightforward formulas for computing:

- Uncertainty (information entropy): $H(Y)$.
- Conditional uncertainty (entropy of Y given knowledge of X): $H(Y|X)$.
- Mutual information (uncertainty reduction): $I(Y; X) = H(Y) - H(Y|X)$.

Step 4: Compare

Results provide quantitative measures of uncertainty and leverage.

We can understand the information effects of the independent variables in terms of the reduction in uncertainty. $H(Y)$ measures how uncertain we are that any observation will have a successful outcome. For each factor, $H(Y|X)$, measures how uncertain we are about whether the case will be a success given that we know whether the factor is present or not. $I(Y; X)$ then measures mutual information—reduction in this uncertainty or information gain—from knowing each factor. When mutual information is close to zero, it means that the theorized factor tells us nearly nothing about the likely outcome. When mutual information is close to the original uncertainty value, it means that the factor almost perfectly predicts the outcome. The magnitude of different factors' mutual information results provides a numerical scale for their systematic comparison indicating their relative contribution to outcome uncertainty reduction.

The numerical results (see Table 7 for an example) enable ranking the qualitative case factors and evaluating their relative as well as cumulative impact. This method thereby enhances qualitative case studies by offering a systematic assessment of the variables' relative impact as well as information gain from factors the case studies deem theoretically important.

We demonstrate the method at work and its analytic leverage with a re-examination of the prominent case study examples discussed above.

Information Analytics and the Limits of Coercive Diplomacy

Returning to Alexander George's landmark work in *The Limits of Coercive Diplomacy* (1971, 1994), we can see a first concrete example of the potential for information theory to enrich the method of structured, focused comparison.

George and his coauthors examine seven cases ($n = 7$), of which three are successful outcomes ($Y = 1$) and four are unsuccessful or ambiguous ($Y = 0$). This is our dependent variable. The uncertainty, $H(Y)$, is 0.985, which reflects the near-even split between successful and other outcomes. Perfect uncertainty—a fifty-fifty split in the outcomes—would be $H(Y) = 1$. (If the outcomes were all successes or all failures, there would be no uncertainty in predicting the outcome: $H(Y) = 0$). We next analyze each factor's impact in an attempt to reduce this uncertainty. Counting the occurrence of each of the factor values relative to the values of the outcome allows us to calculate the joint and conditional probabilities. Table 4 displays the information analytics for each of the factors

TABLE 4. Information in *The Limits of Coercive Diplomacy*

X_i	*Conditional Entropy:* Uncertainty of Y Given X $H(Y\vert X_i)$	*Mutual Information:* Uncertainty Reduction in Y Due to X $I(Y;X_i) = H(Y) - H(Y\vert X_i)$
Clarity of objective	0.96	0.02
Strong motivation	0.99	0.00
Asymmetry of motivation	**0.52**	**0.47**
Sense of urgency	0.86	0.13
Strong leadership	0.86	0.13
Domestic support	0.86	0.13
International support	0.86	0.13
Fear of unacceptable escalation	0.86	0.13
Clarity of terms	**0.52**	**0.47**

(*Notes*. The variables making the strongest contribution to reducing uncertainty are in bold).

identified in *The Limits of Coercive Diplomacy* (Table 7 in the appendix shows all of the calculations).

Table 4 shows that in the coercive diplomacy study, the mutual information values fall into three ranges. First, the *Asymmetry of Motivation* and the *Clarity of Terms* are the only variables that provide predictive leverage with the mutual information of 0.47. This uncertainty reduction suggests that the knowledge of either of these two factors can account for nearly 50% of the expected outcome. Second, the *Clarity of Objective* and *Strong Motivation* tell us essentially nothing about whether or not coercive diplomacy is likely to work (mutual information is nearly zero). In the case of *Strong Motivation*, the absence of information gain is clear because strong motivation is present in every case: it cannot help us discriminate between success and failure. The *Clarity of Objective* is also non-informative, but in this instance that is because it is present in two of the successful cases and two of the failure cases and is absent in two failures and one success. Again, it provides no information to discriminate between successes and failures. Third, for most of the other variables (the *Sense of Urgency*, *Strong Leadership*, *Domestic Support*, *International Support*, and *Fear of Unacceptable Escalation*), the mutual information measure is very low at 0.13 relative to the total uncertainty to be reduced. Each of these variables shows a different pattern, but the shared bottom line is that they do not match up systematically with the pattern of successes and failures.

The findings inform interpretation of results with objective, comparable measures. Information analytics found two most informative variables, each of which can remove nearly half of the uncertainty about the expected outcome. This suggests a relatively large information gain from the two variables—in the face of many sources of uncertainty and many possible causal variables involved in international diplomacy. The contribution of the other variables is relatively scant. The information-theoretic method enables systematically sorting out these degrees of knowledge gain with metrics to inform qualitative interpretations.

The complex subject matter—and its policy relevance—warrants a further investigation of the residual uncertainty. *The Limits of Coercive Diplomacy* illustrates how this can be achieved by leveraging information-theoretic results to help us understand the impact of individual cases. For example, both of the variables found to be most informative have values that line up with all but one of the cases—that is, in the cases of Laos and Cuba, where the outcome is positive, the values for each of these variables are also positive and vice versa. Importantly, it is the Nicaragua case that does not line up for both factors. This suggests a need to look at this case more closely with a particular focus on these two variables.

The Nicaragua chapter in *The Limits of Coercive Diplomacy* illustrates the qualitative ambiguities in underlying case studies (the raw data for the information analysis) that may account for the uncertainty still unexplained. The Nicaragua case discusses the outcome as part of "explaining the limits of success" (George et al. 1994:188). It presents the coercive diplomacy objectives in this case as removal from power of the Sandinista government in Nicaragua, and grants that this did indeed happen in the February 1990 elections, but only after years of intensive US coercive diplomacy efforts. (Hence, the case study authors' use of the term "success" and the positive factual outcome warrant the positive coding of this outcome as "Y" in Table 1 and "1" in the entropy analysis calculations). The chapter also discusses possible limits on the extent of coercive diplomacy's contribution to this outcome in the context of other forces that may have contributed to the outcome, questioning the causal factors.

From a comparative perspective, there are only three successes in this set of cases, and these issues with one of them (Nicaragua) may contribute to lack of information gain from the seven out of nine factors analyzed across all cases. Figure 1 demonstrates the potential for information analysis to help us understand the impact of individual cases on the analysis. It shows the change in the mutual information score when each case is individually dropped from the analysis. As expected, the Nicaragua case has a particularly important effect on the *Asymmetry of Motivation* and *Clarity of Terms* cases. Without the Nicaragua case, both of these would nearly perfectly discriminate between successes and failures. A compete tabulation of both outcomes and case factors as suggested here, plus the information analysis conducted with these tabulated data, provides a clearer view of the results as well as challenges that, made explicit by this analysis, may be anticipated and addressed in future such studies.

Information analysis suggests case selection guidelines to improve comparative results. The selection of a combination of empirically well-documented cases and variables amenable to clear binary classification covering all possible combinations of factors and outcome values facilitates the knowledge gain from qualitative data. When such thorough documentation is unavailable, information analytics still offer leverage to enhance case comparison.

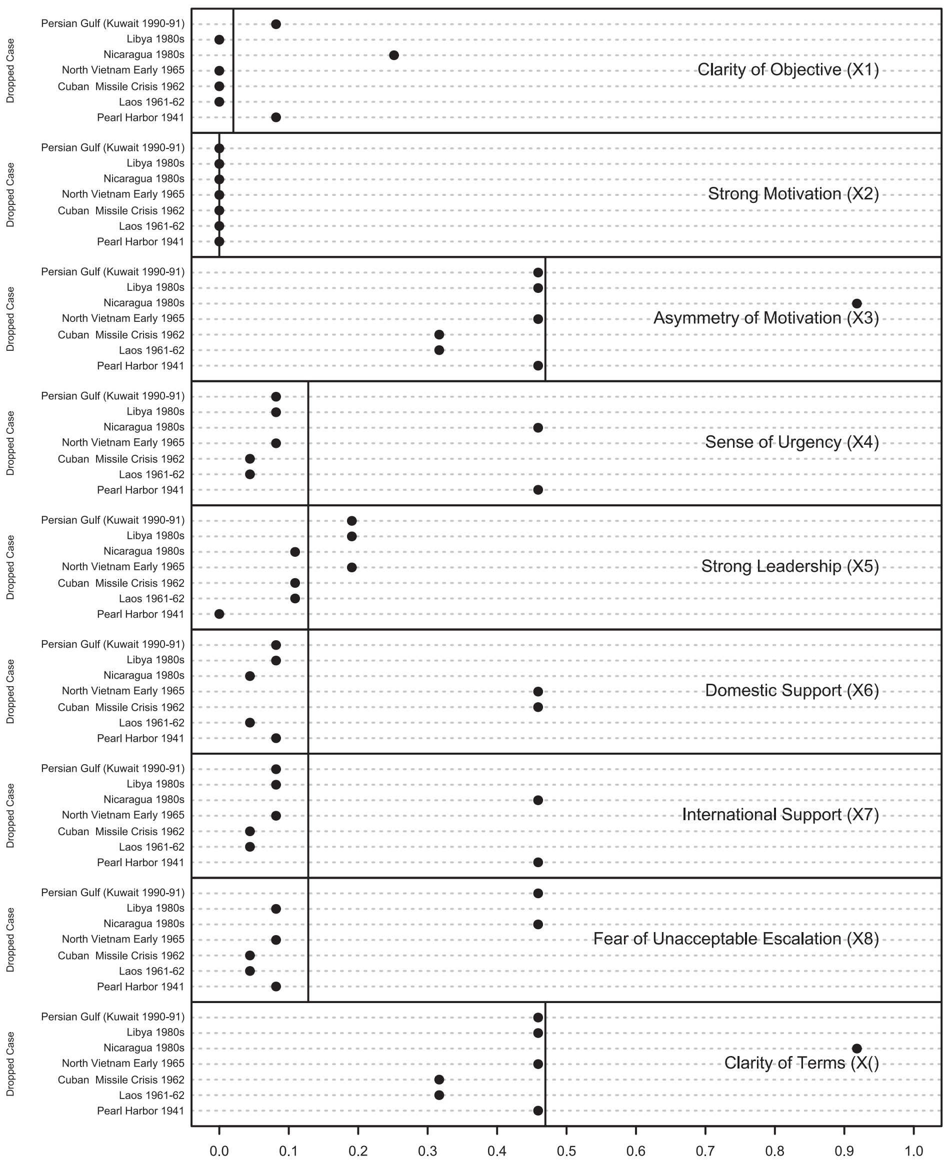

FIG 1. Assessment of Individual Case Effects. This Figure Shows the Impact of Dropping Each Case on the Mutual Information Score for Each Factor. Dots Represent the Information Score Without the Respective Case. The Vertical Lines Show the Aggregate Scores with all Cases Included

Information Analytics and the Politics of Arms Control Treaties

Krepon and Caldwell's *The Politics of Arms Control Treaty Ratification* (1991) also has seven cases, but now with four successes and three failures. In this example, two variables align perfectly with the outcomes: *Presidential skill in handling executive-congressional relations* and the *Support of Senate leadership and pivotal senators*. The presence of positive values for either of these variables perfectly predicts treaty ratification. Table 5 presents the information analytics.

The outcome uncertainty here is exactly the same as in the coercive diplomacy study, $H(Y) = .985$. In both studies, there are seven cases with a three-to-four split between successes and failures. Five of the variables provide minimal discriminatory information, $I(Y;\ X_i) \leq .20$. Two are in the middle range: *Presidential popularity*, ($I(Y;\ X_i) = .47$) and the *Perception of the president as a defender of U.S. national security interests*, $I(Y;\ X_i) = .52$. The two variables with the greatest mutual information nearly equal to the outcome uncertainty, $I(Y;\ X_i) = .98$, are most informative about the outcome (Table 5).

Thus, contrary to the stated concerns of Krepon and Caldwell (1991:400), information analysis allows us if not to completely rank order, then at least to sort the variables into several bins in terms of their impact on arms control treaty ratification, with a clear quantitative measure of the relative magnitude of that impact. As with the George and Simons study of coercive diplomacy (1994), information analysis provided a more explicit understanding of the relative impact, or the lack thereof, for each of the policy-relevant factors.

The precision of information-analytic results benefits from including more cases and as many combinations of outcome and factor states as possible. This dynamic is exemplified in our final case reanalysis.

Information Analytics and the Evolution of UN Peacekeeping

William Durch uses the structured, focused method to compare twenty cases of the UN peacekeeping experiences between 1945 and 1992 on fourteen different factors (1993:12). Durch sets out to test three hypotheses:

Table 5. Information in *The Politics of Arms Control Treaty Ratification*

X_i	*Conditional Entropy:* Uncertainty of Y Given X $H(Y\|X_i)$	*Mutual Information:* Uncertainty Reduction in Y Due to X $I(Y;X_i) = H(Y) - H(Y\|X_i)$
Perception of substantive treaty benefits	0.79	0.20
Presidential popularity	0.52	0.47
Perception of president as defender of U.S. National Security Interests	0.46	0.52
Perception of president as experienced in foreign affairs	0.86	0.13
Presidential skill in handling executive-congressional relations	**0.00**	**0.98**
Quality of presidential advice	0.79	0.20
Favorable international environment	0.79	0.20
Support of senate leadership and pivotal senators	**0.00**	**0.98**
Support of military leadership	0.79	0.20

(*Notes*. The variables making the strongest contribution to reducing uncertainty are in bold).

that successful peacekeeping requires local consent, that it needs great power support, and that there needs to have been some alteration in the basic objectives of the warring parties such that they are ready for a compromise solution. Oddly, as noted above, while the questions around which the analysis is structured are effectively repeated for each case, there is no tabulation of factors by case and no summary of the results. The analysis is, however, sufficiently clear to allow us to build a summary table as shown in Table 3.

Just creating a structured table that includes an outcome variable already significantly improves our ability to visually compare the results of the Durch peacekeeping study. But as the table expands, visual assessment of the different variables becomes increasingly difficult and imprecise.

Simple information analysis provides clarity and precision. Table 6 summarizes the information analytics based on the data in Table 3. In addition to the discipline that information analysis forces on the results presentation, it also provides a way to assess the relative impact of each variable. In Table 6, we see that *Local Consent* and *Attitude Change*, which are assessments of the critical variables for Durch's first and third hypotheses, have the most power for discriminating between success and failure in UN peacekeeping missions. But even these variables only reduce uncertainty by about 50% (uncertainty reduction is 0.506 compared to the underlying uncertainty of 0.971). Additionally, these two variables take the same value in every case, so it is impossible to assess their relative or unique impacts.

Every other variable's uncertainty reduction is near zero. In particular, *Great Power Support* and *US Support*, which accord with Durch's second hypothesis, both show relatively little impact on Peacekeeping success. Instead, Table 3 suggests that American or great power support might be necessary for having a mission at all. Every case has either American or great power support with the singular exception of the UN mission for the referendum in Western Sahara (MINURSO).

Of course, the lack of strong relationships can often be as interesting as their presence. The fact that ten of the

Table 6. Information in *The Evolution of UN Peacekeeping*

X_i	*Conditional Entropy:* Uncertainty of Y Given X $H(Y\|X_i)$	*Mutual Information:* Uncertainty Reduction in Y Due to X $I(Y;X_i) = H(Y) - H(Y\|X_i)$
Local consent	**0.465**	**0.506**
Great power support	0.826	0.144
United States support	0.826	0.144
Attitude change	**0.465**	**0.506**
Domestic conflict	0.96	0.011
International conflict	0.902	0.069
UN Security Council Initiative	0.941	0.03
Local initiative	0.941	0.03
Third-party broker	0.895	0.076
Political mission	0.911	0.06
Military mission	0.968	0.003
Narrow mandate	0.958	0.013
Broad mandate	0.97	0.001
Mandate revision	0.951	0.02

(*Notes*. The variables making the strongest contribution to reducing uncertainty are in bold).

fourteen variables selected for examination in this extensive study fail to provide significant information about the success of peacekeeping missions is itself telling. The context and control variables—the political, military, and geographic setting, the source of UN involvement, the scope of political support for the operation, the nature of the UN mandate, the sources of funding, and the operational problems all show no real relationship to the success or failure of UN peacekeeping missions. This pattern, which is made clear by information analytics, may in turn point to opportunities to gain more insight through reducing possibly "redundant" factors or combining sets of variables via Quantitative Case Analysis (QCA) (Ragin 1987; Rihoux and Ragin 2009). QCA would be complementary, but the advantage of the information-theoretic method is to clarify the original study as is, in its original configuration, plus identify the non-informative variables that may be candidates for reduction via QCA or another such method that alters the original study design in search of further insight.

Disentangling the impact of local consent and attitude change would require additional cases that are distinct on the values of these factors. Information analysis cannot distinguish between these two co-occurring variables based on the case data available from the Durch study. It does, however, bring them more clearly to our attention than originally presented by Durch, thus helping to evaluate the results of this study as well as highlighting opportunities for further study.

Limitations and Extensions

Information theory offers a useful tool for bridging the quantitative/qualitative divide in multi-method analysis. As we have emphasized, however, information analysis is not a panacea for the inherent limits of small-n data. Case and factor selection still needs to be rigorously theory driven rather than ad hoc. Analysts still must be careful about not over-interpreting small differences in the quantitative measures. These are, of course, issues for all qualitative and quantitative methods. Information analysis helps make these issues more explicit and can provide some metrics for better understanding the impact of specific assumptions.

The information-theoretic approach makes assumptions about random variables and independent data, which need to be addressed when analyzing specific cases. The structured, focused samples (sets of cases and factors studied by each underlying case study) need to satisfy these assumptions through theoretically appropriate construction. This can be challenging, but it is a challenge shared more generally by all case study methods, including the structured, focused method (George and Bennett 2005). Information analysis reminds us of the importance of these underlying assumptions by making them more explicit. In the case of structured, focused comparisons, the random assumption is satisfied when any particular case that meets the theoretic selection criteria may in principle be selected. The data independence assumption is satisfied by the nature of what makes the phenomena studied different enough to qualify as distinct cases (for example, different geographic locations, timeframes, sets of participants). It remains important, though, to remember that any limitations in the data selection process, will limit the ability to generalize from the data. The entropy measure is a straightforward metric for describing the relationship between factors and outcomes in a set of cases. As we have presented it here, it is not a procedure for drawing statistical inferences for populations from the characteristics of a random sample.[4] As with any case analysis, one must still be careful about how general inferences are drawn from the specific set of analyzed cases.

An additional limitation that we have seen in the examples provided above, but most clearly in the Durch study, is that the entropy measure cannot discriminate among the effects of factors that vary perfectly together. Again, there is no shortcut for getting around the inherent limits in the data and the limited degrees of freedom of a small number of observations. Still, the application of information-analytic techniques helps to clearly highlight the structure of variation in the data set and will increase our awareness of these limitations. Qualitative case analysis (QCA) can be a helpful follow-on technique for further analyzing the collinearity in a small-n data set.

This brings us to a final point about putting information analysis in the context of the larger universe of small-n analytic techniques. As we have shown here, information analysis can be an accessible and sufficient tool for making structured focused comparison more rigorous and systematic. It can also serve as a gateway to some more complex quantitative approaches to small-n case analysis. Here, we might point to tools, such as qualitative case analysis (Ragin 1987; Rihoux and Ragin 2009), two-by-two causation tests (Braumoeller and Goertz 2000; Seawright 2002), or Bayesian inference models (Western and Jackman 1994; Dion 1998). Each of these methods shares with our approach the goals of advancing rigorous small-n and multi-method research but pursues different purposes and users. Each of these methods makes a valuable contribution and offers other useful approaches to configuring our studies to gain additional insights, for example, through combining or reducing the factors of interest, or incorporating researcher beliefs as prior information for analyzing small-n results. Information analytics, however, help enhance the qualitative studies as originally designed and carried out by their authors without reconfiguration. And, importantly, it is sufficiently straightforward to be accessible to a wide range of researchers.

Conclusion

The movement toward multi-methods is built on the complementary advantages of qualitative and quantitative methods. It starts from the recognition that both modes of analysis draw on the same logic of counterfactual understanding (Fearon 1991) and scientific method (King et al. 1994; Van Evera 1997; Bueno de Mesquita 2002). Usually the multi-method approach has simply combined the systematic precision of a large-N overview, with the depth and nuance of qualitative analysis. Our argument here has been that information analytics can go a step further by giving us some user-friendly tools for the more qualitatively inclined to do systematic assessment of comparative case studies.

Of course, information analysis is not a short cut around the basic problems of too few cases, too many variables, missing data, or biased research design. Nor

[4] Confidence intervals can be drawn for entropy metrics, but with small numbers of observations, and especially if there are any missing data in the underlying case studies, the benefits are rather meager relative to the increase in conceptual and mathematical complexity (Sveshnikov 1978:288; Esteban and Morales 1995).

can it correct for the myriad dangers in conceptualization, operationalization, and measurement that lurk in both quantitative and qualitative methods. What it can do is provide a systematic and consistent way to understand the impact of the independent variables and to assess the empirical results of small-n studies of complex phenomena and policy decisions under uncertainty of many interacting factors and limited information.

This approach also provides tools for better research design. For instance, the analysis identifies just how important some of the theorized variables may be empirically and how they compare to others. Findings may then support more informed research design decisions about which variables and cases to examine in future studies. The most informative factors found may serve as dimensions for exploring contrasting types of cases and conditions toward theory building.

Finally, for policy analysis and design, this approach can help practitioners objectively determine and prioritize the information they have based on uncertainty reduction metrics. The structured, focused comparison method provides powerful tools for problems with a smaller number of cases that require more in-depth analysis. It applies the basic logic of the scientific method to qualitative analysis. With the appropriate attention to the theoretic derivation of variables and careful case selection, it offers probative value. As we have seen in the examples provided here, however, the results can still be ambiguous and dependent on particular cases. Information analytics do not change the number or configuration of variables, nor the amount of information present in a structured, focused comparison; rather, they clarify both the results and the impact of specific factors and cases as conceived by the original studies.

Information analysis is useful to strengthen case study research, simple to calculate based on counts, and integrative—combining quantitative rigor with qualitative depth toward accumulating knowledge and improving understanding. Scholars can use the information-theoretic approach presented here to help evaluate prior research, design new studies, and inform the policy implications of their work. This previously underappreciated approach should have a more prominent place in the political scientist's multi-methods toolbox.

References

Braumoeller, Bear F., and Gary Goertz. (2000) The Methodology of Necessary Conditions. *American Journal of Political Science* 44: 844–858.

Bueno de Mesquita, Bruce. (2002) Domestic Politics and International Relations. *International Studies Quarterly* 46: 1–9.

Carlsnaes, Walter. (1992) The Agency-Structure Problem in Foreign Policy Analysis. *International Studies Quarterly* 36: 245–270.

Congleton, Roger D. (2001) Rational Ignorance, Rational Voter Expectations, and Public Policy: A Discrete Informational Foundation for Fiscal Illusion. *Public Choice* 107: 35–64.

Converse, Philip E. (1985) Power and the Monopoly of Information. *American Political Science Review* 79: 1–9.

Coppedge, Michael. (1999) Thickening Thin Concepts and Theories: Combining Large N and Small in Comparative Politics. *Comparative Politics* 31: 465–476.

Cover, Thomas M., and Joy A. Thomas. (2006) *Elements of Information Theory*, 2nd edn. New York, NY: John Wiley & Sons.

Dhar, Vasant III, Dashin Chou, and Foster Provost. (2000) Discovering Interesting Patterns for Investment Decision Making with GLOWER: A Genetic Learner Overlaid with Entropy Reduction. *Data Mining and Knowledge Discovery* 4: 251–280.

Dion, Douglas. (1998) Evidence and Inference in the Comparative Case Study. *Comparative Politics* 30: 127–145.

Dodd, Lawrence C. (1974) Party Coalitions in Multiparty Parliaments: A Game-Theoretic Analysis. *American Political Science Review* 68: 1093–1117.

Drozdova, Ekaterina [Katya]. (2008) Organizations, Technology, and Network Risks: How and Why Organizations Use Technology to Counter or Cloak Their Human Network Vulnerabilities. Unpublished Doctoral Dissertation, Information, Operations, and Management Sciences, New York University.

Drozdova, Katya, and Michael Samoilov. (2010) Predictive Analysis of Concealed Social Network Activities Based on Communication Technology Choices: Early-Warning Detection of Attack Signals from Terrorist Organizations. *Computational and Mathematical Organization Theory* 16: 61–88.

Durch, William J. (1993) *The Evolution of UN Peacekeeping: Case Studies and Comparative Analysis*. New York, NY: St. Martin's Press.

Eckstein, Harry. (1975) Case Study and Theory in Political Science. In *The Handbook of Political Science*, edited by Fred I. Greenstein and Nelson W. Polsby. Reading, MA: Addison-Wesley.

Esteban, Maria Dolores, and Domingo Morales. (1995) A Summary on Entropy Statistics. *Kybernetika* 31: 337–346.

Fearon, James D. (1991) Counterfactuals and Hypothesis Testing in Political Science. *World Politics* 43: 169–195.

George, Alexander L., and Andrew Bennett. (2005) *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

George, Alexander L., David K. Hall, and William E. Simons. (1971) *The Limits of Coercive Diplomacy: Laos, Cuba, Vietnam*. Boston, MA: Little Brown.

George, Alexander L., and William E. Simons (eds.) (1994) *The Limits of Coercive Diplomacy*, 2nd edn. Boulder: Westview Press.

Gleick, James. (2011) *The Information: A History, A Theory, A Flood*. New York, NY: Pantheon Books.

Guizzo, Erico Marui. (2003) The Essential Message: Claude Shannon and the Making of Information Theory. Unpublished Thesis, Program in Writing and Humanistic Studies, Massachusetts Institute of Technology.

Hogg, Robert V., and Allen T. Craig. (1995) *Introduction to Mathematical Statistics*. New York, NY: Prentice-Hall.

Horgan, John. (1990) Claude E. Shannon: Unicyclist, Juggler and Father of Information Theory. *Scientific American* 242: 20–22.

King, Gary, Robert O. Keohane, and Sidney Verba. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Krepon, Michael, and Dan Caldwell. (1991) *The Politics of Arms Control Treaty Ratification*. New York, NY: Palgrave Macmillan.

Lieberman, Evan S. (2005) Nested Analysis as a Mixed-Method Strategy for Comparative Research. *American Political Science Review* 99: 435–452.

Lowry, Dennis T., and Theodore J. Marr. (1975) Clozentropy as a Measure of International Communication Comprehension. *Public Opinion Quarterly* 39: 301–312.

McCombs, Maxwell, and Jian-Hua Zhu. (1995) Capacity, Diversity, and Volatility of the Public Agenda: Trends from 1954 to 1994. *Public Opinion Quarterly* 59: 495–525.

Midlarsky, Manus I. (1974) Power, Uncertainty, and the Onset of International Violence. *Journal of Conflict Resolution* 18: 395–431.

Molinar, Juan. (1991) Counting the Number of Parties: An Alternative Index. *American Political Science Review* 85: 1383–1391.

Oppenheim, Felix E. (1978) "Power" Revisited. *Journal of Politics* 40: 589–608.

Provost, Foster, and Tom Fawcett. (2001) Robust Classification for Imprecise Environments. *Machine Learning* 42: 203–231.

Ragin, Charles C. (1987) *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Rapoport, Anatol. (1974) War and Peace. *Annals of the American Academy of Political and Social Science* 412: 152–162.

Rihoux, Benoît, and Charles C. Ragin. (2009) *Configurational Comparative Methods: Qualitative Comparative Analysis (QCA) and Related Techniques*. Applied Social Research Methods Series. Thousand Oaks, CA: Sage.

Samoilov, Michael Sergey. (1997) Reconstruction and Functional Analysis of General Chemical Reactions and Reaction Networks. Unpublished Doctoral Dissertation, Biophysics, Stanford University.

Samoilov, Michael, Adam Arkin, and John Ross. (2001) On the Deduction of Chemical Reaction Pathways from Measurements of Time Series of Concentrations. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 11: 108–114.

Seawright, Jason. (2002) Testing for Necessary and/or Sufficient Causation: Which Cases Are Relevant? *Political Analysis* 10: 178–193.

Shannon, Claude E. (1948) A Mathematical Theory of Communication. *Bell System Technical Journal* 27: 379–423.

Shannon, Claude E., and Warren Weaver. (1949) *The Mathematical Theory of Communication*. Urbana: University of Illinois Press.

Sheingate, Adam D. (2006) Structure and Opportunity: Committee Jurisdiction and Issue Attention in Congress. *American Journal of Political Science* 50: 844–859.

Sveshnikov, Aram A. (1978) *Problems in Probability Theory, Mathematical Statistics and Theory of Random Functions*. New York, NY: Dover Publications.

Van Evera, Stephen. (1997) *Guide to Methods for Students of Political Science*. Ithaca, NY: Cornell University Press.

Verdú, Sergio. (1998) Fifty Years of Shannon Theory. *IEEE Transactions on Information Theory* 44: 2057–2078.

Western, Bruce, and Simon Jackman. (1994) Bayesian Inference for Comparative Research. *American Political Science Review* 88: 412–423.

Appendix 1: The Methodology

This appendix provides a brief background on the relevant basic probability concepts followed by practical instructions and equations for implementing the information-theoretic analysis.

Probability Background

Information analysis requires the ability to work with a limited set of probability functions. All probabilities in this analysis are calculated simply based on frequencies of event or factor occurrences (Sveshnikov 1978; Hogg and Craig 1995; Cover and Thomas 2006):

1. The probability of an outcome or a factor (variable) is the number of positive occurrences divided by the total number of observations (cases). This is a frequency estimation of probability. The probability of each factor is written $p(x)$, and the probability of the outcome is written $p(y)$.
2. The joint probability of a particular outcome and a particular factor (variable) combination is the number of their co-occurrences (for example, the number of times both the outcome and the factor are coded 1) divided by the total number of cases. Joint probability is written $p(x, y)$.
3. The conditional probability of a particular outcome given a particular factor (variable) is the probability that y is true given that x is true and is written as $p(y|x)$. The conditional probability is calculated by dividing the joint probability of x and y by the probability of x:

$$p(y|x) = p(x, y)/p(x) \tag{1}$$

Table 7 provides an example of the counts and all subsequent calculations used in the coercive diplomacy study to calculate these probabilities.

Computing Outcome Uncertainty (Information Entropy)

Information entropy is a measure of uncertainty in a variable. For our outcome variable Y, information entropy is (Cover and Thomas 2006: Chapter 2):

$$\begin{aligned} H(Y) &= -\sum_y p(y)\log_2 p(y) \\ &= -p(y=0)\log_2 p(y=0) \\ &\quad -p(y=1)\log_2 p(y=1) \end{aligned} \tag{2}$$

Maximum information entropy occurs at $p(y = 1) = .5$. That corresponds with the point at which we have the

Table 7. Coercive Diplomacy Information-Theoretic Analysis

		Coercive Diplomacy Factors (Independent Variables)								
	Computation	Clarity of Objective	Strong Motivation	Asymmetry of Motivation	Sense of Urgency	Strong Leadership	Domestic Support	International Support	Fear of Escalation	Clarity of Terms
Data frequency counts: (x, y)	count $(x_i = 0, y = 0)$	2	0	4	3	1	1	3	3	4
	count $(x_i = 1, y = 0)$	2	4	0	1	3	3	1	1	0
	count $(x_i = 0, y = 1)$	1	0	1	1	0	2	1	1	1
	count $(x_i = 1, y = 1)$	2	3	2	2	3	1	2	2	2
Joint probabilities: p(x, y)	count $(x_i = 0, y = 0)/$n	0.29	0.00	0.57	0.43	0.14	0.14	0.43	0.43	0.57
	count $(x_i = 1, y = 0)/$n	0.29	0.57	0.00	0.14	0.43	0.43	0.14	0.14	0.00
	count $(x_i = 0, y = 1)/$n	0.14	0.00	0.14	0.14	0.00	0.29	0.14	0.14	0.14
	count $(x_i = 1, y = 1)/$n	0.29	0.43	0.29	0.29	0.43	0.14	0.29	0.29	0.29
Probabilities: $p(x)$	count $(x_i = 0)/$n	0.43	0.00	0.71	0.57	0.14	0.43	0.57	0.57	0.71
	count $(x_i = 1)/$n	0.57	1.00	0.29	0.43	0.86	0.57	0.43	0.43	0.29
Conditional probabilities: $p(y\|x) = p(x, y)/p(x)$	count $(x_i = 0, y = 0)/$ count $(x_i = 0)$	0.67	*	0.80	0.75	1.00	0.33	0.75	0.75	0.80
	count $(x_i = 1, y = 0)/$ count $(x_i = 1)$	0.50	0.57	0.00	0.33	0.50	0.75	0.33	0.33	0.00
	count $(x_i = 0, y = 1)/$ count $(x_i = 0)$	0.33	*	0.20	0.25	0.00	0.67	0.25	0.25	0.20
	count $(x_i = 1, y = 1)/$ count $(x_i = 1)$	0.50	0.43	1.00	0.67	0.50	0.25	0.67	0.67	1.00
Conditional entropy	$H(Y\|X)$	0.96	0.99	0.52	0.86	0.86	0.86	0.86	0.86	0.52
Mutual information	$I(Y;X) = H(Y) - H(Y\|X)$	0.02	0.00	0.47	0.13	0.13	0.13	0.13	0.13	0.47

(*Notes*. *Undefined because there are no cases where $x_i = 0$ for this variable).

maximum uncertainty whether y will be 1 or 0 (that is, there is a fifty-fifty chance of observing either 1 or 0). The more certain we are that y will be either 1 or 0, the lower the information entropy. [If $p(y = 1) = 1$, there is zero uncertainty in the value of y—it is always 1—and, hence, information entropy of y is zero].

Computing Uncertainty Due to Each Variable (Conditional Entropy)

The conditional entropy tells us the uncertainty in outcome Y given knowledge about variable X (Cover and Thomas 2006):

$$\begin{aligned} H(Y|X) &= \sum_x p(x)H(Y|X=x) \\ &= -\sum_x p(x) \quad \sum_y p(y|x)log_2 p(y|x) \\ &= -p(x=0) \\ &\quad *[p(y=0|x=0)log_2 p(y=0|x=0) \\ &\quad +p(y=1|x=0)log_2 p(y=1|x=0)] \\ &\quad -p(x=1) \\ &\quad *[p(y=0|x=1)log_2 p(y=0|x=1) \\ &\quad +p(y=1|x=1)log_2 p(y=1|x=1)] \end{aligned} \quad (3$$

Computing Uncertainty Reduction (Mutual Information)

The mutual information measures uncertainty reduction in outcome Y due to the knowledge of X—or, alternatively, how much information about Y is gained by learning X. In our setting, this means how much information is gained about a policy outcome by learning the value of factors as a result of case studies. The mutual information is calculated by subtracting the conditional entropy for that variable from the total information entropy (Shannon 1948; Shannon and Weaver 1949; Cover and Thomas 2006):

$$I(X;Y) = H(Y) - H(Y|X) \quad (4)$$