Political Methodology

Committee on Concepts and Methods

Working Paper Series

25

February 2010

Structured, Focused Uncertainty

Information Analysis

for Multi-Method Comparative Case Studies

Katya Drozdova

Empirical Studies of Conflict Project (drozdova@stanford.edu)

&

Kurt Taylor Gaubatz

Old Dominion University (kgaubatz@odu.edu)

C&M
The Committee on Concepts and Methods
www.concepts-methods.org

IPSA
International Political Science Association
www.ipsa.org

CIDE
Research and Teaching in the Social Sciences
www.cide.edu

The C&M working paper series are published by the Committee on Concepts and Methods (C&M), the Research Committee No. 1 of the International Political Science Association (IPSA), hosted at CIDE in Mexico City. C&M working papers are meant to share work in progress in a timely way before formal publication. Authors bear full responsibility for the content of their contributions.

The Committee on Concepts and Methods (C&M) promotes conceptual and methodological discussion in political science. It provides a forum of debate between methodological schools who otherwise tend to conduct their deliberations on separate tables. It publishes two series of working papers: "Political Concepts" and "Political Methodology."

Political Concepts contains work of excellence on political concepts and political language. It seeks to include innovative contributions to concept analysis, language usage, concept operationalization, and measurement.

Political Methodology contains work of excellence on methods and methodology in the study of politics. It invites innovative work on fundamental questions of research design, the construction and evaluation of empirical evidence, theory building and theory testing. The series welcomes, and hopes to foster, contributions that cut across conventional methodological divides, as between quantitative and qualitative methods, or between interpretative and observational approaches.

Submissions. All papers are subject to review by either a member of the Editorial Board or an external reviewer. Only English-language papers can be admitted. Authors interested in including their work in the C&M Series may seek initial endorsement by one editorial board member. Alternatively, they may send their paper to wps[at]concepts-methods.org.

The C&M webpage offers full access to past working papers.

www.concepts-methods.org

Acknowledgements

The authors would like to thank Dr. Michael Samoilov, of UC Berkeley's California Institute for Quantitative Biosciences (QB3), for helpful comments and suggestions. Kathlyn Taylor Gaubatz provided helpful editorial feedback. K. Drozdova also would like to gratefully acknowledge the Hoover Institution for the opportunity to serve as a visiting scholar and benefit from research in the Hoover Institution Library and Archives while writing this article.

Introduction

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 approach to enhance small-n case study research.

The method of "structured, focused comparison" was advanced by Alexander L. George (1979) as a response to the heightened interest in more systematic applications of qualitative methods to derive empirical generalizations (Eckstein, 1975). In this approach a set of theoretically motivated critical variables is identified and then their variation is analyzed across several detailed qualitative cases studies to derive systematic conclusions. While the method of structured, focused comparison is admirably systematic and analytic, we argue that its real world applications have often failed to provide analytic clarity. Drawing on recent advances in the field of information theory, we propose a relatively straightforward method to provide a systematic quantitative understanding of the strengths and limitations of structured, focused comparisons.

Traditional statistics are not of much help in assessing comparative case studies. It is not only the small number of observations that challenge meaningful application of traditional statistics to cross-case meta-analysis. A deeper issue involves the nature of complex, nonlinear real-world phenomena marked by the uncertainty of many interacting factors that the in-depth case studies often explore. Typical correlation-based statistics are not well equipped to tackle such problems. Correlations provide good estimates for linear relationships with Normal distributions, whereas in case studies we often deal with more complex systems and networked sets of policy or other socio-political processes and events with unknown or not Normal distributions. The information-theoretic approach is designed to reduce uncertainty and sort out the impacts in such complex and multi-factor settings—common to real world problems studied by political scientists—using reasonably intuitive and simple calculations.

In this paper, we begin with a brief review of the method of structured, focused comparison. Next we present the information theory approach. We demonstrate the analytic process by applying this approach to two prominent examples of structured, focused comparisons. In each example, we show how the information theory approach can sharpen the analysis by providing quantitative measures of the relative impact of multiple variables on the policy outcomes and by estimating their uncertainty levels. The findings enhance our understanding of comparative results and their policy implications as well as generate suggestions for improving information analytics in future research.

The Strengths and Limits of Structured, Focused Comparison

The method of structured, focused comparison integrates the advantages of qualitative methods with the systematic analysis 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 historical 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 select a set of relevant cases—these can be either cross-sectional, or can be 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 powerful analytic tool when a set of cases aligns to clarify the impact of one or two central variables. But the complex phenomena investigated by in-depth case studies tend to involve many factors interacting in nonlinear or unknown ways. The simple visual or conceptual comparison of the results lacks the rigor of quantitative methods particularly when dealing with multiple variables and uncertain results. 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 uncertain case results.

Two prominent examples of structured, focused comparisons illustrate these challenges and serve as a useful test-bed for the meta-analytic approach we are proposing. *The Limits of Coercive Diplomacy* by Alexander George and William Simons (1971; 1994) and *The Politics of Arms Control Treaty Ratification* by Michael Krepon and Dan Caldwell (1991) have been cited as exemplars of the structured, focused methodology (George and Bennett, 2005). In both cases, the authors provide an overview of the results but do not actually provide clear guidance on the relative importance of the factors they are studying on case and policy outcomes. Each study provides a table summarizing the presence or absence of factors (independent variables) examined in the cases. The actual analysis of the tables is largely left to the reader's intuition. Table 1, reproduced from *The Limits of Coercive Diplomacy* (with the addition of a "Success" row showing case outcomes, which the original table lacks), suggests the problem. The two unambiguously successful cases—Laos and Cuba—had positive values for most of the operative 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 just a couple of the cases or a subset of the variables, requiring further analytic tools to study such potential dependence.

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 Escalation		+	+		?	?	+
Clarity of Terms	?	+	+				

Source: reproduced from George and Simon, *The Limits of Coercive Diplomacy* (1994, 288) with the addition of "Success" row showing case outcomes as described by the book's case studies. (Notation: "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.)

Krepon and Caldwell's study of arms control treaty ratification is similarly ambiguous. Their results are reproduced here as Table 2 (again, with the addition of a "Ratification" row showing case outcomes, which the original table lacks). Krepon and Caldwell describe 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).

Table 2: Arms Control Treaty Ratification

	Versailles Treaty	Washington Naval Treaties	Geneva Protocol (1926)	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	+	+		+	+	+	+

Source: reproduced 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. (Notation: “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.)

In both of these studies, the results are not presented in a way that allows for clear assessment of the impact of the posited factors on policy outcome nor for assessing the relative impact of the different cases on comparative findings. Notably, these studies do not even present case outcomes (dependent variable values) as part of their comparative results. (Hence, we listed the "Success" and "Ratification" outcomes in Tables 1 and 2 respectively to rectify this missing information critical for a meaningful comparison). These studies reflect the difficulty of assessing the complex interactions between variables and outcomes, with uncertainty about their mutual impact remaining even after substantive case analysis.

There is much to be said for the ability of the human brain to analyze amorphous qualitative information. Our argument here, however, is that this process can be enhanced by also providing systematic quantitative assessments of the results. The information theory approach can leverage the structured, focused case study method by providing systematic, comparable, and replicable measures of influence for the identified factors.

Uncertainty Analysis: An Information Theoretic Approach for Social Sciences

Although structured, focused studies tend to examine well-documented factors in selected cases, much uncertainty remains about relative importance of these factors and the comparative results and explanations. Small-n samples, the complexity of factor interactions, or lack of knowledge of their underlying distributions limit applicability of typical statistical tests.[1] In this context, a key methodological challenge is to gain the most knowledge from the information generated by case studies—while systematically accounting for the variables' relative contributions and

[1] On a more technical note: small samples from normally distributed populations can be analyzed using Student's *t*-distribution, but problems arise for other distributions. Typical statistical analyses use mean and variance (i.e., only the first and second moments of a probability distribution) assuming Normal distribution. If this assumption is incorrect, however, the analysis will be incomplete and potentially misleading (by neglecting other moments). It will not capture the characteristics of skewed, multimodal, etc. distributions. Alternatively, the information-theoretic approach avoids these problems, as information entropy and mutual information (discussed next) capture the entire distribution.

the degree of confidence in the results. The information-theoretic approach addresses this challenge.

This approach uses the related quantities of *information entropy*—a measure of uncertainty—and *mutual information*—a measure of information that one variable contains about another (Shannon, 1948; Shannon and Weaver, 1949; Cover and Thomas, 1991; Cover and Thomas, 2006). The analysis seeks to reduce entropy thereby reducing the uncertainty about an outcome (the dependent variable) based on the knowledge of some of its attributes or independent variables (Samoilov, 1997; Dhar *et al.*, 2000; Provost and Fawcett, 2001; Drozdova, 2008). This analysis uses a probability-based measure of outcome uncertainty and estimates the reduction in uncertainty due to each variable. Event frequencies are used to estimate probabilities, so the necessary probabilities are easily calculated by counting the number of occurrences and co-occurrences of the independent and dependent variables. Ultimately, this approach provides us with a powerful and stochastically rigorous way of understanding and quantifying case comparisons using information theory.

Information theory emerged from the study of communications to answer fundamental questions about information transfer in noisy environments where people might lack the necessary interest or capacity to understand messages. More broadly, the "information entropy" measure that Claude Shannon introduced in his seminal 1948 paper can also serve as a general measure of information, choice and uncertainty. Using the same mathematical formula, Cover and Thomas (2006) define entropy simply as the uncertainty of a random variable.

The entropy measure is based on simple probabilities and a logarithmic functional form as proposed by Shannon (1948). The logarithmic function expresses a smooth curvilinear relationship between the consistency of co-occurrence of an independent variable with the dependent variable. Uncertainty is zero when the independent variable perfectly co-occurs with the dependent variable and rises at a decreasing rate as co-occurrence declines (see Figure 2 in the methodological appendix). The logarithmic function is particularly convenient because: (1) it is more practically useful for dealing with parameters that tend to vary linearly with the logarithm of the number of possibilities, (2) it is more intuitive by allowing linear comparisons with common standards, and (3) it is more tractable mathematically for analyzing entities that involve many possibilities. The unit for measuring information corresponds to the choice of a logarithmic base. For instance, the simple base 2 corresponds to the information units of binary digits or bits. (We use this binary unit to quantify our example studies' variables and results by coding positive occurrences denoted by "Y" and "+" in Tables 1 and 2 as one, and all others as zero. The resulting set of mutually-exclusive binary values then serves as data for the quantitative information analysis.)

Since early applications to communications and cryptography (Shannon and Weaver, 1949), this approach has gained wider usage. Cover and Thomas (2006) show the fundamental contributions that information theory has made to many fields—from statistical physics (thermodynamics' second law that entropy of an isolated system always increases reflecting nature's tendency toward chaos) and computer science (algorithmic or descriptive complexity also known as Kolmogorov complexity describing data compression) to the philosophy of science and statistical inference (instantiating the Occam's Razor principle that the simplest explanation is best) as well as probability and statistics (error exponents for optimal hypothesis testing and estimation). Recent advances have used information-theoretic methods to study complex living systems—including natural (e.g. Samoilov, 1997; Samoilov *et al.*, 2001) and social systems (e.g. Drozdova, 2008).

Despite the increasing use of information analytics across a range of fields, these methods have seen relatively little use in political science. Shannon's original work on information and communication has been appropriately recognized by political scientists—in the introduction to his 1984 Presidential Address to the American Political Science Association. Philip Converse called it a "watershed study" (1985). Most political scientists have drawn on information theory as a way of understanding political information and communication processes (e.g. 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 her study of a broader sample of organizations to reduce uncertainty about—and gain new insights into—organizations' mission-critical technology strategies based on their mission, network structure, environment, and resources. In that study, the information analysis determined two jointly most informative variables which served as dimensions to select contrasting cases for in-depth comparative case studies of the relationship between technology and security strategies in hostile environments. Here we build on that approach to generalize the use of information theory as a meta-analytic for improving qualitative case study methods.

Uncertainty Reduction Method

This section provides a conceptual description of the basic method and steps for conducting the information entropy and mutual information analysis to sharpen structured, focused case study results. A methodological appendix briefly reviews the basic probability concepts employed by this analysis and provides the equations and further details for the necessary computations.

Information entropy of a random variable Y is denoted by H(Y). In the structured, focused case study setting, Y represents total uncertainty about the outcome (dependent variable). The independent variables analyzed by the cases (denoted by X_i) examine conditions associated with policy outcomes. To determine how informative each X_i may be about an observed set of case outcomes, we calculate conditional entropy denoted by $H(Y|X_i)$.

Conditional entropy measures the information entropy of one random variable given another random variable. The reduction in uncertainty in Y due to our knowledge of X_i is provided by the mutual information denoted by $I(X_i;Y)$. The mutual information is the relative entropy between the joint distribution and the product distribution of two random variables (Shannon, 1948; Cover and Thomas, 2006). In our analysis, it measures dependence between X_i and Y. The mutual information is also the difference between the total and conditional entropy written generally as $I(X;Y)=H(Y)-H(Y|X)$.[2] Mutual information is symmetric, nonnegative and is measured by the same unit as entropy (Shannon, 1948; Cover and Thomas, 2006), i.e., bits in our analysis.

[2] Mutual information can be calculated conditional on several variables as well as sequentially or in terms of other combinations of variables. However, for the example small-n analyses here, we start with a simple binary calculation.

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 variables analyzed. The magnitude of mutual information compared to the original outcome entropy (uncertainty) indicates the magnitude of the uncertainty reduction due to the knowledge of this variable factor. This method supports the ranking and evaluation of the variables' relative importance and contribution to reducing overall outcome uncertainty in the given problem space.

The information-theoretic approach makes assumptions about random variables and independent data, which need to be addressed when analyzing specific cases. (Data in this analysis again are the binary values mutually-exclusively assigned to the factors qualitatively studied in each case). The structured, focused samples support these assumptions about the underlying inputs since, in theoretically constructed samples (George and Bennett, 2005), any particular case—which satisfies the theoretical selection criteria—may in principle be selected. Independence of the quantified inputs, for the purposes of information theoretic study, may be supported by the differences of the analyzed cases which typically involve different sets of participants and situations (e.g., geographical locations, time, historical and political context, etc.).

This mutual information analysis supports a better understanding of results produced by the existing studies based on their qualitative data. Such data themselves may be incomplete, noisy or ambiguous as exemplified by Tables 1 and 2. (That is, there may be some uncertainty about the occurrence of the factors studied.) In such situations, one may further examine the likely range of the mutual information result given levels of confidence in the probability estimate of factor occurrence. For this purpose, an error bound estimation approach (also discussed in the methodological appendix) provides a way to investigate potential tradeoffs between the levels of certainty in the factor presence (probability) and the likelihood of learning new information about the outcome based on this factor. By varying the confidence level on the factor probability, one can explore the level at which the lower bound of the mutual information becomes nonzero—indicating that the results are not random and the uncertainty reduction does occur. Figure 1 summarizes the information analytic process, and the appendix details the practical steps and equations to implement it. The fairly straightforward computation can be managed using an Excel spreadsheet or another such tool.

Overall the information-theoretic approach enhances qualitative case studies by offering a systematic assessment of the variables' relative impact as well as information gain from factors deemed theoretically important by case studies.

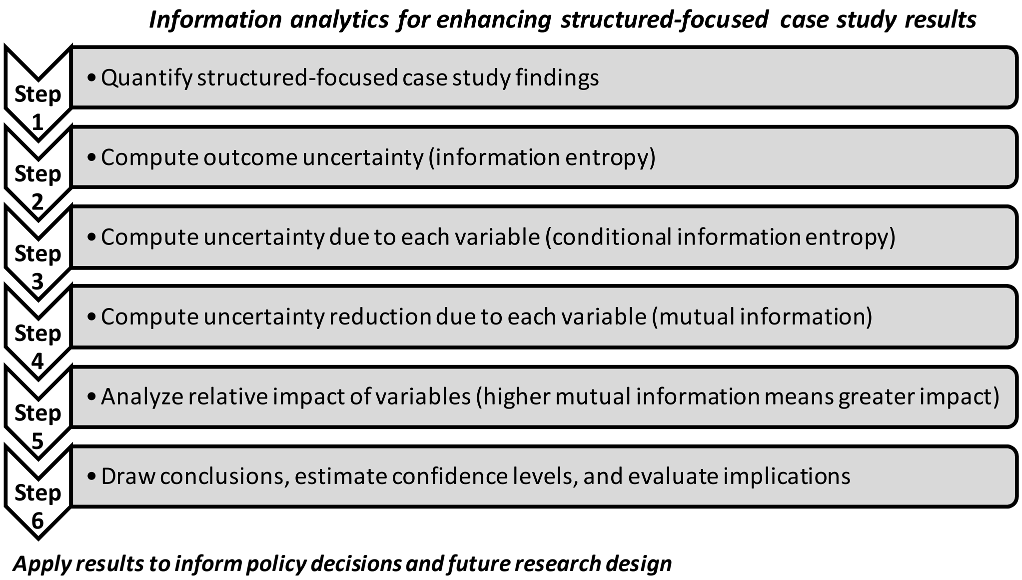

Figure 1: Information analytic steps and process for enhancing structured, focused comparative case studies

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 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). The total outcome uncertainty or information entropy, H(Y), is .985. This is our dependent variable. We next analyze each variable's impact in an attempt to reduce this uncertainty. Counting the occurrence of each of the variable values relative to the values of the dependent variable allows us to identify the joint and conditional probabilities as outlined in the previous section. Table 3 displays the information analytics for each of the factors identified in *The Limits of Coercive Diplomacy* (and Table 5 in the appendix shows the calculations).

Table 3: Information in The Limits of Coercive Diplomacy

	Conditional Info Entropy: **Conditional Entropy of Y given X**	***Mutual Information:*** **Uncertainty Reduction in Y due to X**
X_i	$H(Y\|X_i)$	$I(Y;X_i)=H(Y)-H(Y\|X_i)$
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 Escalation	0.86	0.13
Clarity of Terms	0.52	**0.47**

As we have seen, H(Y|X) is the measure of how uncertain we are about whether the case will be a success (Y=1) given that we know the value of each X.

When H(Y|X) is close to one it means that the causal factor tells us very little about the likely value of Y. When it is close to zero it means that X almost perfectly predicts the outcome. We can see in Table 3 that in the coercive diplomacy study these values fall in three ranges. The *Clarity of Objective* and *Strong Motivation* tell us essentially nothing about whether or not coercive diplomacy is likely to work (H(Y|X) is very close to 1). In the case of *Strong Motivation*, the absence of real information 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. For most of the other variables (the *Sense of Urgency*, *Strong Leadership*, *Domestic Support*, *International Support*, and *Fear of Unacceptable Escalation*) the entropy measure is still very high at .86 relative to the total entropy. For each of these variables there is a different pattern, but the bottom line is always that they do not line up well with the pattern of successes and failures.

The *Asymmetry of Motivation* and the *Clarity of Terms* are the only variables that provide much predictive leverage with entropy measures of .52. We can understand the information effects of the causal variables in terms of the reduction in uncertainty. The outcome uncertainty, H(Y), measures how uncertain we are that any observation will have a successful outcome. In the diplomacy case, with three successes out of seven cases, this uncertainty is very high (H(Y) = .985). We can then calculate mutual information, i.e., the reduction in uncertainty due to any of the posited causal factors, by subtracting the conditional uncertainty, H(Y|X), from the outcome uncertainty, H(Y). In the coercive diplomacy analysis, this gives us a reduction in uncertainty of .47 for the *Asymmetry of Motivations* and the *Clarity of Terms*. This reduction in the overall uncertainty suggests that knowledge of these two factors can account for nearly 50 percent of the outcome.

Interpretations of this finding may vary based on the study's goals. The information findings inform interpretation with objective measures. Specifically, the information analytics found two most informative variables, each of which can remove nearly half of the uncertainty about the expected outcome. This suggest a relatively large information gain from the two variables—in the face of many sources of uncertainty and change involved in the complex matters of international diplomacy. Additionally, the findings identify the relatively scant contribution of the other variables. The information-theoretic method enables us to systematically sort out these degrees of information or knowledge gain and provides assessment metrics to inform qualitative interpretations.

The inherently challenging and complex nature of the subject matter—as well as its policy relevance, where case study findings may be used to inform future policies and actions—warrant a further investigation of possible residual sources of uncertainty. *The Limits of Coercive Diplomacy* again serves well to illustrate how this can be achieved by further 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.

Additionally, the Nicaragua chapter discusses the outcome as part of "explaining the limits of success" (George and Simons, 1994, 188). It presents the coercive diplomacy objectives in this case as removal from power of the Nicaraguan Sandinista government, and grants that this government was indeed removed from power in the February 1990 elections following years of US coercive diplomacy efforts. (Hence the case study authors' use of the term "success" and the positive factual outcome warrants the positive coding of this outcome as "Y" in Table 1 and "1" in the entropy analysis calculations). However, 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. Furthermore, from a comparative perspective, there are only three successes in this set of cases, and these issues with one of them (Nicaragua) may present non-trivial limitations on the overall analysis. This suggests possible sources of error due to uncertainty about the underlying case factors whose probabilistic range can be estimated further.

In this context, to evaluate the likely range and potential limitations of the results, we use the error bounds analysis (detailed in the methodological appendix). Using a 95% confidence level on the factor probability estimate, we find identical nonzero error bounds on the outcome uncertainty reduction due to *Asymmetry of Motivation* and the *Clarity of Terms*. For each of these variables, the mutual information result ranges from a minimum of .02 to maximum of .77 bits at 95% confidence in the probability of occurrence of each of these factors. That is, with 95% confidence in factor occurrence, knowledge of either of these factors does reduce our uncertainty about the outcome.

With such a small number of cases, the confidence bounds *per se* may be less important than the nonzero lower bound's implication. It allows us to conclude that the results of George and Simons' (1994) structured, focused case comparison with respect to *Asymmetry of Motivation* and the *Clarity of Terms* are not merely random and we do learn about coercive diplomacy outcomes from knowledge of these conditions. The mutual information finding indicates the extent of this learning. For each of these factors, the .47 mutual information—in the system with .985 total uncertainty—means that knowledge of each of these two factors can reduce nearly by half the uncertainty about the coercive diplomacy policy outcome. Though specific qualitative interpretations may vary, this finding provides a quantitative measure of uncertainty reduction and indicates the relative importance of these factors to coercive diplomacy results. The other factors include zero in their lower error bound, which brings into question their ability to significantly reduce outcome uncertainty.

An additional lesson learned from this analysis is that for best results in terms of determining the degree of learning from the variables it is best to select a combination of cases and variables that cover all possible combinations of factors' occurrence and outcome states.

Information Analytics and the Politics of Arms Control Treaties

Krepon and Caldwell's *The Politics of Arms Control Treaty Ratification* (1991) provides another opportunity to examine the use of information analysis for enhancing our understanding of structured, focused comparisons. In this analysis there are, again, seven cases, but now with four successes and three failures. In this example, however, we have two variables that 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. The information analytics are presented in Table 4 (and calculations in Table 6 in the appendix).

Table 4: Information in The Politics of Arms Control Treaty Ratification

	Conditional Info Entropy: **Conditional Entropy of Y given X**	*Mutual Information:* **Uncertainty Reduction in Y due to X**
	H(Y|X)	I(Y;X)=H(Y)-H(Y|X)
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

The outcome uncertainty in this case is exactly the same as in the coercive diplomacy case (H(Y)=.985). In both examples there are seven cases with a three to four split between successes and failures. Five of the variables provide minimal discriminatory information (H(Y|X)>.75)). Two are in the middle range: presidential popularity (H(Y|X)=.52), and the perception of the president as a defender of U.S. national security interests (H(Y|X)=.46). The two variables that align perfectly with the outcomes are, of course, perfectly discriminatory (H(Y|X)=0). Each of these two variables, *Presidential skill in handling executive congressional relations* and *Support of Senate leadership and pivotal senators*, provides greatest uncertainty reduction about the treaty ratification outcome as shown in Table 4.

The mutual information error bound estimate for these results is inconclusive due to the underlying data limitations—too many missing combinations of outcome and factor states (occurrences or non-occurrences). This again suggests a more general lesson for enhancing structured, focused case studies. To enable a systematic estimation of the range and extent of learning that can be achieved about outcomes from given case factors, the set of cases and factors would benefit from including as many combinations of outcome and factor states as possible.

Overall, 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. 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 Power of Combining Case Methods and Information Analysis

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). Usually this 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 tools for the systematic assessment of the qualitative research component.

Information analysis is not a short cut around the basic problems of too few cases, too many variables, or other data limitations. Nor 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 empirical results in small-n studies of complex, non-linear phenomena.

This approach also provides tools for better research design. For instance, the analysis allows a test of just how important some of the variables—claimed by the literature to be key for understanding particular problems—may be in reality 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.[3] The most informative factors found can also serve as dimensions for exploring contrasting types of cases and conditions with the results contributing to theory building.

Finally, for policy analysis and design, this approach can help practitioners focus on the most important aspects out of several identified in the literature and prior experience. The information-theoretic findings can help guide action choices toward desired policy outcomes. In this way, the information theoretic method for assessing the impact, tradeoffs and courses of action associated with different factors—in complex and uncertain situations where simple linear correlations would not suffice—offers the potential for better informed policy decisions.

Conclusion

The method of structured, focused comparison is a powerful tool for many midrange problems where we have 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 theoretical derivation of variables and careful case selection, it offers significant probative value. As we have seen in the examples provided here, however, the results of structured, focused comparisons can sometimes be ambiguous and dependent on particular cases. Methods drawn from information theory can help us quantify both the results and the impact of specific variables and cases on our analysis.

As we demonstrated in our re-analysis of two prominent examples of structured, focused comparison, the information theory approach sharpens the

[3] Additionally, a technical suggestion for sample selection is to select—while satisfying structured, focused theoretical conditions—a set containing all possible combinations of variable and outcome values for a full problem space analysis. That is, nonzero joint probabilities of each variable and outcome occurrence are preferred, should this be possible empirically, to avoid undefined values and logarithmic blow-up in probabilistic calculations.

analysis and provides additional insights into the comparative results. It offers tools to systematically evaluate the knowledge gained from prior studies and real-world cases. Results may also help guide future research or policy efforts.

The results of applying information theory to structured, focused comparisons may sometimes be disconcerting: the unfortunate reality of small-n analysis is that the outcomes are both probabilistic and highly dependent on particular cases. This is simply the effect of the complexity of underlying phenomena, as well as the inexorable logic of too many variables and too few cases. The use of information theory does not change the amount of information present in a structured, focused comparison. It does, however, help us understand and explicitly communicate the degree of certainty we can derive from the analysis and the sensitivity of the results to specific factors and cases. Importantly, the mathematics of information analysis are relatively straightforward, and can be easily managed with Excel or other simple calculating tools.[4]

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 underappreciated approach should have a more prominent place in the political scientist's multi-methods toolbox.

[4] The methodological appendix provides the Excel spreadsheet calculations for the analyses presented in this paper. These can easily be adapted for other cases.

Methodological Appendix

This appendix provides a brief background on the relevant basic probability concepts followed by practical instructions and equations for implementing the information-theoretic analysis of structured, focused comparative case studies. This step-by-step discussion expands upon the analytic process outline in Figure 1, aiming not only to explain the information-theoretic method but also to guide easy replication of our analyses as well as application of the method to other case studies. Excel is used to enable a simple computational implementation. Table 5 shows the information theoretic calculations for *The Limits of Coercive Diplomacy* and Table 6 for *The Politics of Arms Control Treaty Ratification*.

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 the proportion of specific alignments and combinations of event or factor occurrences (Hogg and Craig 1995, Cover and Tomas 2006, and Sveshnikov 1978 provide a detailed background):

- The ***probability*** of an outcome or a factor (variable) is the number of positive occurrences divided by the total number of cases. This is a frequency estimation of probability. The probability of each case study variable is written p(x) and the probability of outcome is written p(y).
- The ***joint probability*** of a particular outcome and a particular factor (variable) combination is the number of their joint 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 as p(x,y).
- 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:

$$p(y|x) = \frac{p(x,y)}{p(x)} \qquad \text{Equation 1}$$

Tables 5 and 6 exemplify the counts used to calculate these probabilities.

1. Quantifying the structured, focused case study findings to generate data

To generate the data for information-theoretic analysis, an Excel spreadsheet is populated with binary values for each case and factor per qualitative case studies. The factors analyzed by the case studies are independent variables X_i, and case outcome is the dependent variable Y. For each, the value "1" is assigned to a positive occurrence (the presence of a particular factor X or outcome Y) and "0" to others (including negative or ambiguous) to ensure mutually-exclusive coding.

2. Computing outcome uncertainty (information entropy)

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

Equation 2

$$H(X) = -\sum_{x} p(x)\log_2 p(x)$$

$$= -p(x=0)\log_2 p(x=0) - p(x=1)\log_2 p(x=1)$$

Our outcome variable is Y, whose information entropy (uncertainty) is:

Equation 3

$$H(Y) = -\sum_{y} p(y)\log_2 p(y)$$

$$= -p(y=0)\log_2 p(y=0) - p(y=1)\log_2 p(y=1)$$

As shown in Figure 2, the general graph of information entropy in this binary system creates a smooth curve that runs from 0 up to 1 and then back to 0. Maximum information entropy occurs at p(y=1) = .5. That maximum corresponds with the point at which we have the maximum uncertainty whether y will be 1 or 0 (i.e., 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.)

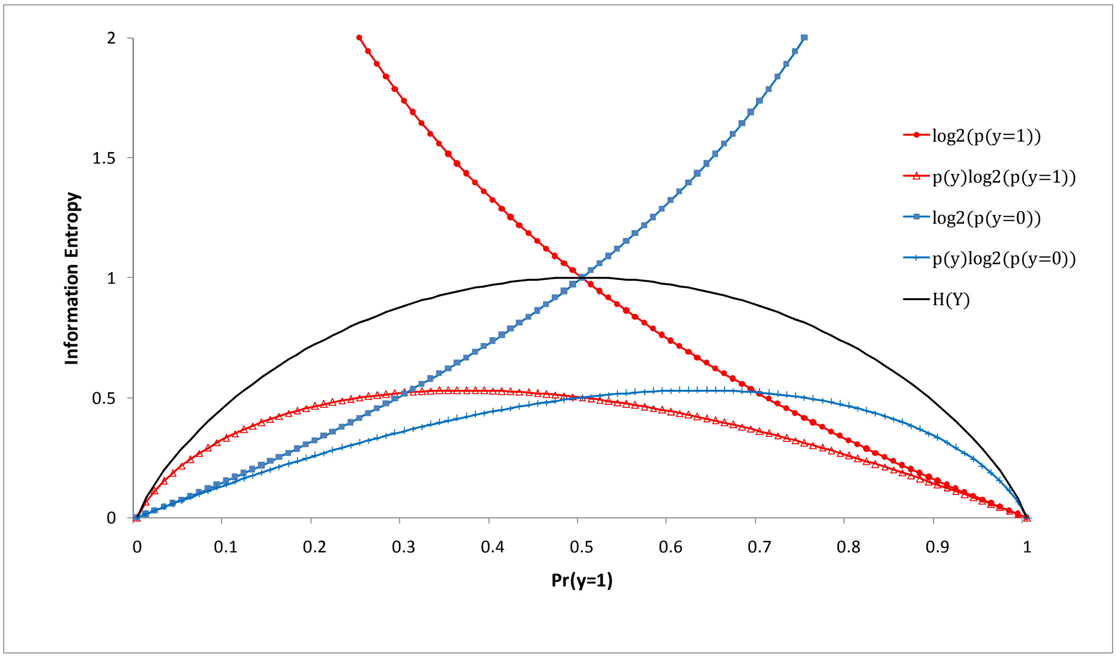

Figure 2: Information entropy in a binary system

3. Computing uncertainty due to each variable (conditional information entropy)

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

Equation 4

$$H(Y|X) = \sum_x p(x) H(Y|X = x)$$

$$= -\sum_x p(x) \sum_y p(y|x) \log_2 p(y|x)$$

$$= -p(x=0) * [p(y=0|x=0)\log_2 p(y=0|x=0) + p(y=1|x=0)\log_2 p(y=1|x=0)]$$

$$-p(x=1) * [p(y=0|x=1)\log_2 p(y=0|x=1) + p(y=1|x=1)\log_2 p(y=1|x=1)]$$

4. Computing uncertainty reduction due to each variable (mutual information)

The mutual information measures uncertainty reduction in outcome Y due to the knowledge of X—or, alternatively, how much certainty (information) about X is gained by knowing or learning Y. The mutual information is calculated by subtracting the conditional information entropy for that variable from the total information entropy (Cover and Thomas, 2006):

Equation 5

$$I(X;Y) = H(Y) - H(Y|X)$$

Mutual information can also be calculated as (Cover and Thomas, 2006):

Equation 6

$$I(X;Y) = H(X) + H(Y) - H(X,Y)$$

where H(X,Y) is joint information entropy of X and Y with a joint distribution p(x,y):

Equation 7

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log_2 p(x,y)$$

$$-p(x=0,y=0)\log_2 p(x=0,y=0)$$

$$-p(x=1,y=0)\log_2 p(x=1,y=0)$$

$$-p(x=0,y=1)\log_2 p(x=0,y=1)$$

$$-p(x=1,y=1)\log_2 p(x=1,y=1)$$

(This will be used to evaluate the relative information gain from each variable as well as confidence in the overall results.)

5. Analyzing variables' relative impact (information gain or uncertainty reduction)

The mutual information results for each variable can now be compared with the outcome entropy to assess the reduction in uncertainty due to each variable. This provides a measure of each variable's importance in terms of providing insight into the uncertain outcome. Furthermore, by comparing the mutual information results of each variable, we can determine their relative impact and importance. This comparative analysis allows ranking or clustering variables according to their impact, identifying most and least informative ones, and so on. Figure 3 illustrates the general relationship between information entropy and mutual information via a Venn diagram.

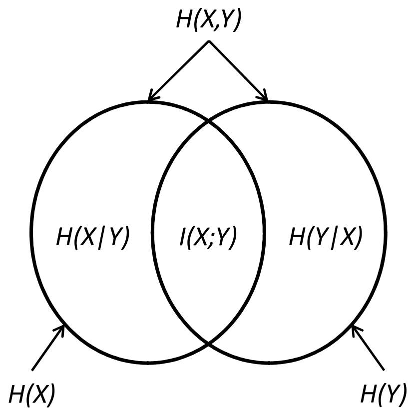

Figure 3: Relationship between entropy and mutual information

Source: reproduced from Cover and Thomas, *Elements of Information Theory* (2006, 22)

6. Estimating confidence bounds to inform conclusions and implications

As part of drawing conclusions and evaluating implications, the researchers may further probe the confidence and range of the information analysis results. For instance, in situations where case data involve some noise or uncertainty, one may conclude that the occurrence of factors (measured by independent variables) varies with some probability. Then, one may examine the likely range of the mutual information result given a level of confidence in the probability estimate.

For these purposes, mutual information is essentially a model of a function that relates case outcomes and conditions, each of which occur with some probability p. There is uncertainty in the mutual information function estimate due to the uncertainty in the estimate of probabilities of events or factors in the case studies. (That is, unless the number of events is infinite, any estimate of the probability for any combination of x and y would be subject to some uncertainty.) By estimating error bounds on the mutual information, we essentially test a null hypothesis that the results may be random. A nonzero lower bound allows the rejection of the null hypothesis and conclusion that new information has indeed been gained from the comparative case study. The range of that information (or knowledge gained) is indicated by the bounds, and the likely magnitude of learning—by the mutual information result itself.

Traditional statistical confidence interval estimation could be used to estimate such uncertainty if mutual information were a linear function of p. Then confidence intervals on p could be simply substituted into the mutual information equation to find its likely range at a chosen confidence interval (CI) level. However, mutual information is a non-linear function, so this approach cannot be applied directly. In this context, one approach is to estimate error bars or likely minimum and maximum bounds on the mutual information results as follows.[5]

The mutual information I(X;Y) expressed as a function of p can be linearized using a first-order Taylor expansion:

Equation 8

$$I(<\vec{p}> \pm \Delta\vec{p}) = I(<\vec{p}>) \pm \sum_i \frac{\partial I}{\partial p_i} \Delta p_i$$

where:

- $I(<\vec{p}>)$ is the mutual information result evaluated by using standard probability estimate (e.g., frequency)—treated here as an estimate whose probabilistic range we evaluate further using the error bounds,
- $\frac{\partial I}{\partial p_i}$ denotes partial derivatives with respect to each of the four possible combinations of x and y (when both are 0, both are 1, one is 1 and the other 0, and vice-versa). Using the conservation condition (that the total probability, i.e. the sum of probabilities of these four states, equals 1), only three partial derivatives will suffice, substituting the fourth, e.g.:

Equation 9

$$p(x=0, y=0) = 1 - p(x=0, y=1) - p(x=1, y=0) - p(x=0, y=0)$$

- Δp_i refers to the confidence interval between the minimal $\mathbf{p}_{\mathbf{min}}$ and maximal $\mathbf{p}_{\mathbf{max}}$ bounds on the underlying probability estimate (of case factor occurrence) calculated at a chosen confidence level (e.g., 95%).

[5] The authors thank Dr. Michael Samoilov, of UC Berkeley's California Institute for Quantitative Biosciences (QB3), for suggesting this error bounds estimation method.

Simplifying:

Equation 10

$$I(\hat{p} \pm \Delta p_i) - I(\hat{p}) = \Delta I(X;Y) = \frac{\partial I(X;Y)}{\partial p_{0,1}}\Delta p_{0,1} + \frac{\partial I(X;Y)}{\partial p_{1,0}}\Delta p_{1,0} + \frac{\partial I(X;Y)}{\partial p_{1,1}}\Delta p_{1,1}$$

Finally, the bounds on Δp can be found using (Sveshnikov 1978, 288):

Equation 11

$$\sum_{j=\mathbf{m}}^{n} \square = \frac{\mathbf{1} - \alpha}{\mathbf{2}}$$

Equation 12

$$\sum_{j=0}^{m} \square = \frac{1 - \alpha}{2}$$

where α is the confidence level for p (e.g., 95%), n is the number of independent observations (i.e., the number of cases) among which the event of interest (i.e., the presence of a factor) occurs exactly m times, and C_n^j is the binomial coefficient.

These equations can be solved for p_{min} and p_{max}, respectively, by using the incomplete Beta function as follows:

Equation 13

$$p_{\min} = \beta^{-1}\left(\frac{1 - \alpha}{2}, m, n - m + 1\right)$$

Equation 14

$$p_{\max} = \beta^{-1}\left(\frac{1 + \alpha}{2}, m + 1, n - m\right)$$

By varying the confidence level on the probability of factor occurrence, the analysis can also explore the relative levels at which the lower-bound on mutual information becomes nonzero for different variables, and how they compare in terms of relative confidence of information gain.

In cases where missing data preclude the calculation of partial derivatives due to logarithmic blow-up (where probabilities are zero), mean values of the interval may be substituted (as long as total probability does not exceed one). An important tradeoff, however, is that while the mean would allow error bar estimation in such cases, the mean biases the probability estimate. Alternatively, our original frequency based approach provides an unbiased estimate of probability, but may preclude straightforward error bar estimation in cases where some of the factor and outcome combinations have not been observed.

Table 5: 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 (xi=0, y=0)	2	0	4	3	1	1	3	3	4
	count (xi=1, y=0)	2	4	0	1	3	3	1	1	0
	count (xi=0, y=1)	1	0	1	1	0	2	1	1	1
	count (xi=1, y=1)	2	3	2	2	3	1	2	2	2
joint probabilities: p(x, y)	count(xi=0, y=0)/n	0.29	0.00	0.57	0.43	0.14	0.14	0.43	0.43	0.57
	count(xi=1, y=0)/n	0.29	0.57	0.00	0.14	0.43	0.43	0.14	0.14	0.00
	count(xi=0, y=1)/n	0.14	0.00	0.14	0.14	0.00	0.29	0.14	0.14	0.14
	count(xi=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(xi=0)/n	0.43	0.00	0.71	0.57	0.14	0.43	0.57	0.57	0.71
	count(xi=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(xi=0, y=0)/ count(xi=0)	0.67	*	0.80	0.75	1.00	0.33	0.75	0.75	0.80
	count(xi=1, y=0)/ count(xi=1)	0.50	0.57	0.00	0.33	0.50	0.75	0.33	0.33	0.00
	count(xi=0, y=1)/ count(xi=0)	0.33	*	0.20	0.25	0.00	0.67	0.25	0.25	0.20
	count(xi=1, y=1)/ count(xi=1)	0.50	0.43	1.00	0.67	0.50	0.25	0.67	0.67	1.00
Conditional Information 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

*=undefined because there are no cases where xi = 0 for this variable

Table 6: Arms Control Treaty Ratification Information-Theoretic Analysis

	Computation	Treaty Ratification Factors (Independent Variables): Perception of 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
data frequency counts: (x,y)	count (xi=0, y=0)	1	2	3	3	3	1	1	3	1
	count (xi=1, y=0)	2	1	0	0	0	2	2	0	2
	count (xi=0, y=1)	0	0	1	3	0	0	0	0	0
	count (xi=1, y=1)	4	4	3	1	4	4	4	4	4
joint probabilities: p(x, y)	count(xi=0, y=0)/n	0.14	0.29	0.43	0.43	0.43	0.14	0.14	0.43	0.14
	count(xi=1, y=0)/n	0.29	0.14	0.00	0.00	0.00	0.29	0.29	0.00	0.29
	count(xi=0, y=1)/n	0.00	0.00	0.14	0.43	0.00	0.00	0.00	0.00	0.00
	count(xi=1, y=1)/n	0.57	0.57	0.43	0.14	0.57	0.57	0.57	0.57	0.57
probabilities: p(x)	count(xi=0)/n	0.14	0.29	0.57	0.86	0.43	0.14	0.14	0.43	0.14
	count(xi=1)/n	0.86	0.71	0.43	0.14	0.57	0.86	0.86	0.57	0.86
conditional probabilities: p(y|x) = p(x,y)/p(x)	count(xi=0, y=0)/ count(xi=0)	1.00	1.00	0.75	0.50	1.00	1.00	1.00	1.00	1.00
	count(xi=1, y=0)/ count(xi=1)	0.33	0.20	0.00	0.00	0.00	0.33	0.33	0.00	0.33
	count(xi=0, y=1)/ count(xi=0)	0.00	0.00	0.25	0.50	0.00	0.00	0.00	0.00	0.00
	count(xi=1, y=1)/ count(xi=1)	0.67	0.80	1.00	1.00	1.00	0.67	0.67	1.00	0.67
Conditional Information Entropy	H(Y|X)	0.79	0.52	0.46	0.86	0.00	0.79	0.79	0.00	0.79
Mutual Information	I(Y;X)=H(Y)-H(Y|X)	0.20	0.47	0.52	0.13	0.99	0.20	0.20	0.99	0.20

References

Carlsnaes, Walter (1992). "The Agency-Structure Problem in Foreign Policy Analysis." *International Studies Quarterly* 36(3): 245-270.

Congleton, Roger D. (2001). "Rational Ignorance, Rational Voter Expectations, and Public Policy: A Discrete Informational Foundation for Fiscal Illusion." *Public Choice* 107(1/2): 35-64.

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

Coppedge, Michael (1999). "Thickening Thin Concepts and Theories: Combining Large N and Small in Comparative Politics." *Comparative Politics* 31(4): 465-476.

Cover, Thomas M. and Joy A. Thomas (1991). *Elements of Information Theory*. NY: John Wiley & Sons.

Cover, Thomas M. and Joy A. Thomas (2006). *Elements of Information Theory*. NY: Wiley-Interscience.

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.

Dodd, Lawrence C. (1974). "Party Coalitions in Multiparty Parliaments: A Game-Theoretic Analysis." *The American Political Science Review* 68(3): 1093-1117.

Drozdova, Ekaterina (Katya) (2008). Organization, Technology, and Network Risks: How and Why Organizations Use Technology to Counter or Cloak Their Human Network Vulnerabilities. New York, New York University. Ph.D.

Eckstein, Harry (1975). "Case Study and Theory in Political Science." In *The Handbook of Political Science*. Fred I. Greenstein and Nelson W. Polsby, Eds. Reading, MA, Addison-Wesley**:** 79-138.

Fearon, James D. (1991). "Counterfactuals and Hypothesis Testing in Political Science." *World Politics* 43(2): 169-95.

George, Alexander L. (1979). "Case Studies and Theory Development: The Method of Structured, Focused Comparison." *Diplomacy: New approaches in history, theory, and policy*: 43–68.

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

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

George, Alexander L. and William E. Simons (1994). *The Limits of Coercive Diplomacy*.

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

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

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

Lieberman, Evan S. (2005). "Nested Analysis as a Mixed-Method Strategy for Comparative Research." *American Political Science Review* 99(03): 435-452.

Lowry, Dennis T. and Theodore J. Marr (1975). "Clozentropy as a Measure of International Communication Comprehension." *The Public Opinion Quarterly* 39(3): 301-312.

McCombs, Maxwell and Jian-Hua Zhu (1995). "Capacity, Diversity, and Volatility of the Public Agenda: Trends from 1954 to 1994." *The Public Opinion Quarterly* 59(4): 495-525.

Midlarsky, Manus I. (1974). "Power, Uncertainty, and the Onset of International Violence." *The Journal of Conflict Resolution* 18(3): 395-431.

Molinar, Juan (1991). "Counting the Number of Parties: An Alternative Index." *The American Political Science Review* 85(4): 1383-1391.

Oppenheim, Felix E. (1978). "`Power' Revisited." *The Journal of Politics* 40(3): 589-608.

Provost, Foster and Tom Fawcett (2001). "Robust Classification for Imprecise Environments." *Machine Learning* 42(3): 203-231.

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

Samoilov, Michael (1997). Reconstruction and Functional Analysis of General Chemical Reactions and Reaction Networks. Stanford, CA, Stanford University. Ph.D.

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.

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(4): 844-859.

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

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