Cumulation from Proper Specification:
Theory, Logic, Research Design, and 'Nice' Laws[1]

Harvey Starr
Department of Political Science
University of South Carolina
Columbia, South Carolina, USA

Introduction

Jim Ray has set out a challenge to students of international conflict: how can we improve our research to do better in the cumulation of knowledge and understanding of international phenomena? While his specific concern has been with the ways in which we go about doing multivariate analysis, the challenge is a much broader one. While most of the commentary in this special issue revolves about more technical statistical issues, the questions that Ray raises must also be addressed not just on statistical/methods grounds, but within the context of broader theoretical concerns and theoretical specification. Simply, that is the aim of this article. Drawing on the concept of the "research triad" of logic, theory, and methodology presented by Most & Starr (1989, 9-17), this article will stress the relationship of theory to the development of research designs that will permit both the additive and integrative cumulation of knowledge. More specifically, this article will discuss how statistical model specification can be <u>informed</u> by <u>broader principles</u> of research design. In so doing, the article will show how Ray's basic points are supported by earlier, complementary, and converging lines of argument; work that did not specifically look at the issues of control or the specific design of multivariate studies, but at the overall process of logic of inquiry and research design.

The primacy of theory is a central theme in Most & Starr, among many others. They call for a number activities. Perhaps first and foremost is understanding the <u>logical form of the</u> <u>relationship</u> generated by a theory (which includes the logical causal structure such as necessity, sufficiency, being necessary and sufficient, etc.), which affects most of the other research design aspects as well– e.g. should we expect linear or curvilinear results, and what is the shape of the relationship between independent and dependent variables? For example, using logic Most &

Starr indicated how scholars had set up research hypotheses based on sufficient relationships, but then went on to test them with designs that could only deal with necessary relationships, and vice versa. Their initial questioning of how researchers have tried to test or evaluate necessary relationships has been taken up more recently in collaborative work by Goertz & Braumoeller (see, e.g., Braumoeller & Goertz, 2003; see also Goertz & Starr, 2003a). One central issue presented in Goertz & Starr (2003b) and Goertz (2003) is that the use of "standard" statistical techniques will not capture the effects of necessity.[2]

Logic, as developed in Most & Starr (1989) also dictates the need to understand more fully what (types of) cases are needed to test or evaluate theory. Case selection is thus intimately related to understanding the context in which the theory should be expected to hold (or, in which one could expect to find evidence of the implications of the theory; see also Achen in this issue). Understanding the broader concepts that are of concern in the theory, as well as understanding what phenomena are really under investigation and why (e.g. Most & Starr, 1989, ch.4 on conceptualizing war) are also needed. Specifying this context involves "nice laws." The broader understanding of concepts and process will also involve "substitutability." Each of these concepts will be addressed more fully below.

As argued in Most & Starr (1989), all of these activities must take place for the proper specification of the theory, which is needed for the accumulation of knowledge. "Proper model specification" is here used in its broadest conceptual and theoretical sense. Nothing I say here will tell any individual scholar exactly how to specify any model in any specific research project. As noted, this is not a technical/methods discussion, directing the reader to specific techniques, but one trying to present broader implications. Here proper specification means asking any

researcher: to think through the nature and logic of the causal relationships involved in the theory being used/developed; to think "process;" to understand the implications of nice laws and substitutability, etc. Indeed, the point here, as in Most & Starr, is to ask the researcher to think through the theoretical structure more fully and clearly, and to include those variables (and in their proper role– e.g. as independent, intervening, complementary or control) that are relevant to the concepts in the theoretical process. This procedure stands in distinction from what Achen in this issue calls "garbage can regressions." Therefore, theoretical specification should be addressed before dealing with the more technical aspects of research design.[3] The explicit argument in Most & Starr (and more implicitly in Ray's papers) is that such activities will have a positive impact on cumulation.

Most & Starr (1989,7), drawing on Zinnes, note that "additive cumulation occurs when 'one study adds some information to the existing literatures on the subject,' through such activities as the citation of previous findings, using previously collected data, secondary or reanalysis of existing data, the incorporation of new cases or new variables into the analysis, or expanding the application of models, indices or techniques to new cases or research questions." Integrative cumulation, however, goes further. It reflects instances where earlier studies are "'crucial' to the conceptual and theoretical components of the subsequent study's research design." Both forms of cumulation need to be pursued, as additive cumulation is an indispensable part of the process that leads to integrative cumulation.

The arguments that Ray presents (2003a, 2003b), can be easily translated to, or subsumed under, this broader enterprise of "cumulation." That is, how can we account for the lack of cumulation of knowledge in certain research areas, and what lessons concerning the logic of

inquiry and research design could we apply to increase both the additive and integrative

cumulation of knowledge in these areas. Most & Starr (1989) similarly used the area of

international conflict to illustrate a set of research design problems that had hindered cumulation

in the study of international relations. They also proposed solutions that had broader

implications. Most & Starr were concerned with the failure to think through the many

complicated aspects of research design, especially the multiple feedback loops among the

components of research design, along with a rather rote application of the basic social science

statistical cookbook. It is possible, then, to see the basic enterprise and general thrust in Ray's

work and in the Most & Starr volume as one and the same– a critical view of how we, as

scholars design our research, and the problems we can fall into when we do not think through the

full ramifications that our research design choices might have for the validity of our results.

Applying the Research Triad: Theory, Nice Laws, and Substitutability

Ray begins with a basic point that he uses to explain the increased (and indiscriminate) use of

control variables. He takes one route to argue the primacy of theory in the selection and use of

control variables, and I will take another. The basic point Ray develops (2003a, 2-4) is that over

the past 15-20 years researchers have moved beyond searching for "ideal" or "best fit" general

models. Instead, they have developed models "to evaluate the impact of one key factor"(2003a,

3), seeing if the impact of that factor can survive to addition of control variables.

Even though its practitioners are almost certainly not thinking in these terms, the latter

"key factor" trend is clearly within the spirit of Most & Starr's idea of "nice laws." These are

"sometimes true 'domain specific' laws" (1989, 98). Most & Starr remind us that while we

should aim for generality, the "right type of law" is one which is clearly specified; that the relationships among variables that it proposes will work only under specified conditions.[4] Most & Starr question whether social scientists will ever generate important "universal" laws. They note, however (1989, 117) that:

> ...it may be useful to recognize that there could very well be laws that are in some sense "good," "domain specific," or "nice" even though the relationships they imply are not necessarily very general empirically... it may be more productive to think of laws each of which is always true under certain conditions (or within certain domains) but which is only "sometimes true" empirically because those conditions do not always hold in the empirical world.

Again, it is likely that "key factor" analysts are <u>not</u> thinking in terms of nice laws, and are still searching for universal laws when they add control variables. Yet, they are on the right track. Even if they do not know what they are doing, this more recent strategy can be viewed as a form of nice laws. It is one way to see under what conditions a key factor does <u>not</u> hold, through the use of control variables. In the spirit of additive cumulation, such studies indeed provides new information through the incorporation of new variables. For example, Ray (2003b, 5), discussing Bremer's work on dangerous dyads, stresses: "Again, the addition of one explanatory factor to the multivariate model originally constructed by Bremer (1992) has several substantial impacts on several of the other relationships that are analyzed. The inclusion of this one additional factor results in a fundamental revision of the rank orderings of the variables by the weight of their apparent causal impact." [5]

The important question is whether or not studies of the impact of some key variable can be part of a process of integrative cumulation. The answer rests with theory, and the place of theory in the research process. Ray supports the "key factor" approach, using a theoretical critique of the ideal/best fit approach (2003a, 4):

> We have no theory, or theoretical approach that will tell us what are the best six, seven, or eight predictor variables to put into a multivariate model aimed at accounting for interstate war or conflict... we have no sound theoretical basis for saying "these are the seven most important variables on which to base any explanation of interstate wars. It is these seven and only these seven..."

What Ray is pointing out is that researchers need theory– they need it first and they need it to be specified. In order for the theory to be specified, that is to understand which factors/variables/etc are important when, under what conditions, and in what form, researchers require a model, and a model which strongly incorporates some theoretical process (e.g.,see Bremer & Cusack, 1996; see also Starr, 1996). In order for theory to be specified researchers require a procedure by which they can build process models. Lave & March (1975) provide such procedures for "conceptual modeling." These procedure include thinking in terms of process as one of their three "rules of thumb." Bremer (1996a, 11) helps us to capture the idea of process by contrasting Mohr's "variance theory" and "process theory": "Variance theories are deterministic, continuous, essentially static theories of being while process theories are stochastic, discrete, inherently dynamic theories of becoming." Bremer's (1996b, 267) description of "process theory" as "a set of interconnected rule-like statements that purport to explain how and under

what conditions a system will undergo change" helps pull together the Lave & March model

building process and Most & Starr's notion of nice laws.

Most & Starr (1989) rely heavily on the Lave & March model building process, including

their own variations such as the use of stylized facts for quick and dirty tests of the implications

of the theory/model. The theoretical specification of a model or theory requires an

understanding of nice laws. That is, the proposed relationships or effects of a model are seen to

hold only under specified conditions– which derive from the process within the model-theory-

story. Most & Starr develop a strategy for honing in on such conditions. Rather than using

elaborate or sophisticated statistical analyses,[6] Most & Starr (1989, 136) look for "stylized facts",

that is: "Any argument that might be developed must be capable of dealing with existing

empirical evidence– stylized facts. While these observations are simple and intentionally

impressionistic, they may not be entirely trivial."

The Most & Starr strategy for taking a first cut at a problem was to develop a process

model, and try to think through "what the world would look like if" one or another of a set of

hypotheses held; hypotheses were not seen as "contending," but rested on nice laws and

proceeded from different conditions. These thought experiments were done rather quickly,

generating rather crude and impressionistic predictions or derivations. Given these directional

indicators, a second step was taken– a 'quick and dirty,' intentionally non-systematic

consideration of some stylized facts and simple hypothetical cases. They note (1989, 182) that

such procedures fit quite closely with Eckstein's (1975, 118) typology of case studies, especially

his notion of 'must fit' cases. Thus, the use of stylized facts shows how case studies could be

used to help delineate nice laws, or the conditions under which different models/theories might

be expected to work.[7] At the other end of the methodological spectrum is a more current approach also concerned with linking theory with methods, and linking theoretical process to empirical analysis– the EITM-related literature (e,g,, see Empirical Implications of Theoretical Models, 2002). As Aldrich & Alt (2003, 309) note, "the problem of separation between good theorizing and good empirical work is common across many disciplines."[8]

The theoretical specification of a model must take into account not only process and nice laws, but also the possibility of "substitutability" (Most & Starr, 1989; see also Cioffi-Revilla & Starr, 1995). Substitutability may be seen as the flip side of nice laws, and may also be approached through the use of stylized facts. Substitutability refers to the existence of a set of alternative modes of response by which decision makers could deal with the same situation. These behaviors have also been called "alternative modes of redundancy" (Cioffi-Revilla &Starr, 1995, 456-57). Most & Starr (1989, ch.5) note that opportunity or willingness can operationally occur or be made available in a <u>number</u> of alternative, non-unique ways. Alternative possibilities or bases for choice produce "substitutability," which Most & Starr see as crucial for understanding the logic of causality, and thus, research design. Morgan & Palmer (2000) use a general meaning of substitutability where any single cause may have multiple effects, and that any single effect may have multiple causes. In the same way, McGinnis (2000) notes that different problems may lead to similar responses, or that a given problem may be dealt with in multiple ways. Bennett & Nordstrom (2000) observe that if policy makers can find multiple paths to success– actions that would all achieve the stated policy goals– then those policies may be thought of as substitutable.

In their investigation of the interaction of opportunity and willingness with substitutability, Cioffi-Revilla & Starr (1995) distinguish between a first-order causality of world politics (opportunity and willingness) at the analytical level, and a deeper second-order causality of substitutability at the operational level. The first order (<u>necessary</u>) elements of opportunity and willingness are linked by the Boolean AND, while the range of possible modes of (<u>sufficient</u>) second order substitutability are connected by the Boolean OR. Cioffi-Revilla & Starr (1995, 453) use their own version of stylized facts when they ask: "What would the world of politics be like if the first-order causality [of opportunity and willingness] were governed by disjunction (OR)?... In such a hypothetical world... behavior could occur when actors had political willingness to act, <u>or</u> when they had the opportunity to do so?" This demonstrates how stylized facts can provide a powerful way of looking at the logical and causal consequences of mis-specifying the form of the relationship. Theory, nice laws, and substitutability will all be used below tin commenting on Ray's "guidelines" for the use of control variables.

Can cumulation occur, and if so, do the relationships between logic, theory, and research design noted above (and inherent in Ray's arguments about the use of control variables) have an effect? Cumulation is a <u>social process</u> that occurs within a community of scholars. One can argue that it has occurred in areas such as the democratic peace, power transition theory, and the role of territory in conflict. Integrative cumulation, for Zinnes (1976) occurs "when earlier studies are 'crucial' to the conceptual and theoretical components of the subsequent study's research design" (Most & Starr, 1989, 7). For example, Russett & Starr (2000) argue that it was only after the actual phenomenon of the democratic peace was empirically established – that two "democracies" do not fight "wars" against each other– that intensive work on theories as to why

this should be occurred. Such theoretical work involved intensive conceptual discussions of

democracy and war, as well as their possible relationships. As theories were generated and

evaluated, they were examined in a variety of different historical and regional contexts. The

implications of these theories, which relied on process, were subsequently tested on a variety of

different behaviors (dependent variables), both conflictual and cooperative. It was then, after a

variety of theories proposed certain types of behaviors based on the nature and presence of

democracy, that researchers expanded their dyadic investigations to monadic studies of

democracies (with Ray being one of the first to argue strongly for such monadic effects).

While Ray uses some examples from articles on the democratic peace for his critiques,

this does not refute my assertion that there has been significant cumulation in the democratic

peace literature. The additive cumulation, as noted, is obvious. The movement from one set of

questions to another, also indicates integrative cumulation. Chernoff (2004), drawing from

discussions of "progress" in the philosophy of science literature, argues that the study of the

democratic peace exhibits the qualities of progressive science and cumulation. One important

indicator, according to Chernoff (2004, 71) is the quality of the "debate" or "dialogues" between

"defenders and opponents of democratic peace hypotheses," demonstrating important aspects of

consensus and progress. In discussing cumulation, it is important to delineate two types of

consensus– it can occur with recognition of research findings as well as in agreement over

theories that explain those results. Chernoff has focused more on the former.

In regard to the latter, my view (based on Zinnes' discussion) is that cumulation does not

demand unanimity or strong consensus on a single theoretical position, but rather on the process

of building towards fuller and deeper findings, and more fully specified theory that has

developed along a variety of branches. It also involves more broadly accepted findings and theory. There is no doubt that this is currently the case regarding the basic findings of democratic peace researchers. But such cumulation also applies to the theory behind the democratic peace. Currently, very few (if any) scholars argue that either the normative/value theory or the structural/constraint theory fully explains the democratic peace to the exclusion of the other. Rather, several related syntheses have been evolving, each based in what Russett and Starr (2000) call the "strategic perspective;" (for instance, the survival/selectorate model of Bueno de Mesquita, et al., 2003).

It is important to note, given the Oneal & Russett article in this issue, that Chernoff cites their work as an important element in the scientific "progress," of the democratic peace research program. Indeed, the Russett & Oneal project has been driven very clearly by theory. It has moved theoretically from very simple models to the more complex model outlined in *Triangulating Peace* (2001, ch.1), using only a very small number of variables whose inclusion and relationship to one another have been theoretically and clearly specified. The spirit and message of the Achen article in this issue is to drop the garbage can approach, where researchers throw in all the variables they think might have an impact (including the use of large numbers of control variables). I see Achen's Rule of Three within this context– not to be rigidly applied, but to force us to think through our theory and what variables we include. Russett & Oneal have done that quite well. With the democratic peace, then, integrative cumulation has been found, but only after hard theorizing, specification of theories, the use of nice laws and substitutability, and substantial additive cumulation have taken place. It has not come through the use of wholesale, general, global models including some set of "best fit" variables.

Illustrations with Ray's Guidelines

Ray (2003a) offers a set of five "guidelines" that he believes could "improve the quality of evidence produced by multivariate analyses, as well as the credibility and intelligibility of that evidence." Let me provide some brief comments applying my points to each of these guidelines.

Guideline #1 states: "do not control for intervening variables." Discussing an article that simply adds control variables "...that previous research has shown to have an impact on war and dispute involvement" (2003a, 5), Ray argues that research design should not treat intervening variables as control variables. The point of my arguments in the present discussion is that if one wants to move from additive cumulation to integrative cumulation, then researchers must recognize exactly <u>what types</u> of variables they are dealing with, and then build a "better" model/theory. If the variables are truly intervening, then the model must be more closely specified to indicate in what ways the new variables are part of the story, and what that means for the implications of the model, and under what conditions the model will hold. A good example is found in the Kadera & Mitchell article in this issue. In a study of one of the implications of theories explaining the democratic peace, Kadera & Mitchell investigate whether international institutions intervene in the relationship between the strength of democratic community and third party conflict management. Rather than just control for joint membership in international organizations, they look at alternative stories– or processes– of how these three variables could be related, and then devise research designs to evaluate these relationships. Kadera & Mitchell, citing Blalock (1964), note that we should not control for a factor that is part of the causality process. Thus, rather than simply throwing what could be an intervening variable

into the hopper of control variables, we must specify (or re-specify) theories, theoretical processes, and look again at the logical form of the relationships implied. Kadera & Mitchell conclude that international organizations do not appear to be an intervening variable, and thus can move on to other considerations.

Similar comments apply to the next three guidelines as well. Specifically, Guideline #2 states: "distinguish between complementary and competing explanatory factors." Ray (2003a,7) notes that in many cases, control variables are simply "alternative causes," and that "rival explanations" are simply alternative causes. One major point developed in Most & Starr, and deriving from nice laws, is the position that very few models or theories are actually "contending;" the key question, proposed in Most & Starr (1989) is rather "how can both be true?" That is, under what conditions does each hold.? Thus, clearer specification and the logic of nice laws helps to support Ray's guideline. One must also investigate whether "alternative" causes are truly "alternative." That is, does substitutability come into play here? Are the alternatives simply substitutable, second order factors by which the behavioral units involved satisfy the first order factors of opportunity or willingness?(see Cioffi-Revilla & Starr 1995).

Similarly, all of these points apply also to Guidelines #3: "do not introduce factors as control variables merely on the grounds that they have an impact on the dependent variable," and #4 as well: "do not control for variables that are related to each other or the key explanatory factor by definition." Especially regarding #4, substitutability can be used here to distinguish between truly substitutable factors or merely multiple measures of the same factor. Again, for #3 and #4 taking nice laws and substitutability into account drives the researcher to more clearly and fully specified theory.

Ray (2003b) discusses the complexity of multivariate analyses more fully. However, Most & Starr reach many of the same conclusions as Ray by discussing the use of quasi-experiments, and looking at different factors as "treatments" which are either present or absent, and the relationship of the treatment to the occurrence of some dependent variable (see Most & Starr, 1980 and Siverson & Starr, 1991 for the use of these ideas in the study of the diffusion of conflict and war). Ray (2003b,11) notes , "Basically, my argument here is that multivariate models of interstate conflict should be simplified." This can be done in many cases using the "treatment" strategy. Nice laws would also ask for simplifying analyses by looking at how relationships stand up to different conditions– using methods as simple as contingency tables! (a strategy noted by Achen in this issue; see also Ray, 2003b, 5).

For example, Ray quotes Achen's (2002, 446), Rule of Three: "a statistical specification with more than three explanatory variables is meaningless." As noted above, Achen critiques the inclusion of long lists of explanatory variables into equations in order to control for their effects, and specifically argues that linear link functions are not "self-justifying." He argues that linear links can make good equations go bad. Such variables should only be included (as I have argued here) for good theoretical reasons. Such variables should be included only after fairly extensive preliminary data analysis reveals what Most & Starr called "the form of the relationship" to the analyst. Indeed, as I have noted above, Achen explicitly notes that preliminary data analysis should include "careful graphical and crosstabular analysis."

While appearing in his discussion of guideline #4, the following example provided by Ray (2003a, 18) illustrates these points:

Running two different analyses, one with contiguity as the control variable for

geographic proximity and another with distance between capital cities as a control

variable, is a perfectly legitimate strategy. Selecting the model with the "best"

results from the point of view of the analyst is also justified., especially if the

analyst can come up with some plausible theoretical conjecture that might

account for the differences in the two analyses. This strategy is clearly preferable,

in my view, to one involving the inclusion of both contiguity and distance as

control variables in the same model or analysis.

Let me waffle here. I fully agree with the general principles in Ray's comments in terms of

how we treat two variables that are similar or measure the same concept. The specific example

noted, however, is useful in that it shows how researchers can differ in their interpretations of

concepts and indicators. Oneal & Russett in this issue have responded by indicating what they

see as differences in the concepts of proximity and distance from that of contiguity, and how the

measures may tap different things. While I might disagree with their specific positions (waffle

again), the key point is that Oneal & Russett have presented reasonable theoretical/conceptual

arguments for using both variables. Ray's commentary is reasonable in questioning a design that

(in his interpretation) uses two indicators of the same phenomenon.

Guideline #5 states: "control for possible differences between across space and over time

relationships." Within the spirit of nice laws, #5 looks at the most frequently used, and perhaps

most important of those conditions that would vary, and affect the applicability of our

models/theories. Time and space are key elements in many studies that look to see "under what

conditions" a model or theory might hold. Starr (2003) argues that, generally, time has taken

precedence over spatial factors (and for a variety of understandable reasons). Most of our research has stressed the temporal dimension or temporal context. This is easily seen in the design of our research where temporal patterns are central, the use of time series data and designs are standard, and the use of time to delineate of our units of analysis is standard. Scholars routinely divide time into historical eras or different international "systems" with a temporal boundary. We do so because we think that theoretically (either explicitly or implicitly) some conditions have changed so much, or on some important dimension, that these changed conditions should alter the way that humans behave or how our theories should work (or cease to work). Changes in technology, especially, as well as the changes in geopolitical structure that often follow wars, also affect crucial space-time relationships (e.g. Boulding's loss-of-strength gradient). Simply put, guideline #5 is extraordinarily important, especially from a nice law perspective.

   These five guidelines for the use of control variables are presented by Ray as a response to one statistical problem that occurs in the research on international conflict. However, the problems that he identifies with the improper use of control variables represent broader issues in the relationship between theory, logic, and research design. Ray (2003a, 28) points to the broader issue of cumulation in his conclusion:

> Were these guidelines to be adopted in future research efforts, the results of empirical evaluations of multivariate models might be more readily interpretable and understood. Ideally, research on international conflict would then be more cumulative, as well as more productive of valuable insights into those processes leading to interstate conflict and war.

The present article has also attempted to indicate how the issue of control variables can be tied to a more general concern with cumulation. In so doing I have tried to indicate the utility of returning to a number of concepts and strategies to improve the cumulative research enterprise initially raised by Most & Starr (1989).

References

Achen, C. 2002. Toward a new political methodology: Microfoundations and ART. *Annual Review of Political Science* 5: 423-450.

Aldrich, J., and J. Alt. 2003. Introduction to the Special Issue. *Political Analysis* 11(4): 309-315.

Bennett, D., and T. Nordstrom. 2000. Foreign policy substitutability and internal economic problems in enduring rivalries. *Journal of Conflict Resolution* 44(1): 33-61.

Braumoeller, B.F. 2003. Causal complexity and the study of politics. *Political Analysis* 11(3): 209-233.

Braumoeller, B. F., and G. Goertz. 2003. The statistical methodology of necessary conditions. In *Necessary conditions: theory, methodology, and applications*, eds. G. Goertz and H. Starr. Lanham, MD: Rowman & Littlefield.

Bremer, S. A. 1992. Dangerous dyads: conditions affecting the likelihood of interstate war, 1816-1965. *Journal of Conflict Resolution* 36 (2):309-341.

Bremer, S.A. 1996a. Advancing the scientific study of war. In *The process of war: Advancing the scientific study of war,* eds. S. A. Bremer and T. R. Cusack. Philadelphia: Gordon and Breach.

Bremer, S.A. 1996b. Final words. In *The process of war: Advancing the scientific study of war,* eds. S. A. Bremer and T. R. Cusack. Philadelphia: Gordon and Breach.

Bremer, S.A., and T. R. Cusack, eds. 1996. *The process of war*. Philadelphia: Gordon and Breach.

Bueno de Mesquita, B., et al. 2003. *The logic of political survival*. Cambridge, MA: MIT Press.

Chernoff, F. 2004. The Study of democratic peace and progress in international relations. *International Studies Review* 6 (1): 49-77.

Cioffi-Revilla, C., and H. Starr. 1995. Opportunity, willingness, and political uncertainty: Theoretical foundations of politics. *Journal of Theoretical Politics* 7(4): 447-476

Eckstein, H.S. 1975. Case study and Theory in political science. In *Handbook of political science, vol.7, Strategies of inquiry*, eds. F. Greenstein and N. Polsby. Reading, MA: Addison-Wesley.

Empirical Implications of Theoretical Models Report. (2002). EITM Workshop, July 9-10, 2001. (Available from: http://www.nsf.gov/sbe/ses/polisci).

Goertz, G. 2003. Cause, correlation, and necessary conditions. In *Necessary conditions: theory, methodology, and applications*, eds. G. Goertz and H. Starr. Lanham, MD: Rowman & Littlefield.

Goertz, G. 2005. Necessary condition hypotheses as deterministic or probabilistic: Does it matter? *Qualitative Methods* 3(1):22-27.

Goertz, G., and H. Starr, eds. 2003a. *Necessary conditions: theory, methodology, and applications*. Lanham, MD: Rowman & Littlefield.

Goertz, G., and H. Starr. 2003b. Introduction: Necessary condition logics, research design, and theory. In *Necessary conditions: theory, methodology, and applications*, eds. G. Goertz and H. Starr. Lanham, MD: Rowman & Littlefield.

Holsti, K.J. 1991. *Peace and war: Armed conflicts and international order, 1648-1989*. Cambridge: Cambridge University Press.

Lave, C. A., and J. G. March. 1975. *An introduction to models in the social sciences*. New York:
Harper and Row.

McGinnis M. D. 2000. Policy substitutability in complex humanitarian emergencies: a model of
individual choice and international response. *Journal of Conflict Resolution* 44(1): 62-89.

Morgan, T. C., and G. Palmer. 2000. A model of foreign policy substitutability. *Journal of
Conflict Resolution* 44(1): 11-32.

Most, B.A., and H. Starr. 1980. Diffusion, reinforcement, geopolitics, and the spread of war.
American Political Science Review 74(4): 932-946.

Most, B. A., and H. Starr. 1989. *Inquiry, logic and international politics*. Columbia, SC:
University of South Carolina Press.

Palmer, G. ed. 2000. Special Issue on "Substitutability in Foreign Policy: Applications and
Advances." *Journal of Conflict Resolution* 44 (1).

Ray, J. L. 2003a. Explaining interstate conflict and war: What should be controlled for? *Conflict
Management and Peace Science* 20(1): 1-31.

Ray, J. L. 2003b. Constructing multivariate analyses (of dangerous dyads)" Paper presented at the
2003 Peace Science Society (International) meeting, Ann Arbor, November.

Russett, B., and J. Oneal. 2001. *Triangulating peace*. New York: W.W. Norton.

Russett, B., and H. Starr. 2000. From Democratic peace to Kantian peace: Democracy and
conflict in the international system. In *Handbook of war studies II,* ed. M.I. Midlarsky.
Ann Arbor: University of Michigan Press.

Siverson, R. M., and H. Starr. 1991. *The diffusion of war: A study of opportunity and willingness*.
Ann Arbor: University of Michigan Press.

Starr, H. 1996. Advancing the scientific study of war: A commentary. In *The process of war: Advancing the scientific study of war,* eds. S. A. Bremer and T. R. Cusack. Philadelphia: Gordon and Breach.

Starr, H. 2003. The power of place and the future of spatial analysis in the study of conflict. *Conflict Management and Peace Science* 20(1): 1-20.

Zinnes, D. A. 1976. The problem of cumulation. In *In search of global patterns,* ed. J.N. Rosenau. New York: Free Press.

Endnotes

1. I would like to thank Kelly Kadera and Sara McLaughlin Mitchell, as well as two anonymous referees for their comments on earlier drafts of this article, which was originally presented at the annual meeting of the Peace Science Society (International), November 2004, Houston. The responsibility for all errors and omissions, of course, is my own.

2. This is not to imply that sufficiency is any more easily captured, only that there are alternatives to what Goertz (2003) calls "correlational hypotheses." See also Goertz (2005), as well as the responses to his piece in the "Necessary Conditions" Symposium in *Political Analysis.*.

3. For how more technical methodological choices may follow, see Most & Starr (1989, chs. 3-4) for the effects of logic and different forms of causality on case selection, and the analyses of those cases; see the chapters in Goertz & Starr, 2003) for the methodological impact of "necessary condition hypotheses," and designs structured around dealing with necessity as the form of the

relationship; or Cioffi-Revilla and Starr (1995) who use various logics to deal with the uncertainty of politics. Discussions of specific methods such as standard interaction terms, or Boolean methods, should be left to true methodologists– such as Braumoeller (2003).

4. For example, a theory of third-party deterrence which holds in the 19th century might not hold under conditions which include nuclear warfare. The earlier debate over the effects of bipolarity vs. multipolarity essentially argued that under one condition the balance of power (stability) would work well, but not under the other. Ray (2003b, 4), for example, discusses the different results of Bremer's research on the relationship between development and war in terms of different conditions, in that the structure and complexity of the models used were quite different. In physics almost all "universal" laws provide conditions that must be met for the "law" to hold.

5. Reflecting a point made below on substitutability, one question that must be asked is whether such results truly reflect the impact of nice laws, or simply statistical inadequacy in the face of complexity.

6. Again, let me make clear that the concept of "nice laws' does not presuppose any specific context, set of conditions, or form of relationship. The concept only asks the researcher to think about and hypothesize when and where– what spatial and temporal domains– and under what special conditions (or "treatments" as in the diffusion studies of Most & Starr, 1980 and Siverson & Starr, 1991) the causal relationships should hold. Holsti (1991) provides a good example of the temporal dimension, in his wide ranging discussion how the issues that lead states to war have changed over the centuries. The treatments/conditions involved might be very simple (e.g.

bordering a state which is at war in regard to the spatial contagion of conflict, or the existence of ICBM delivery systems with nuclear warheads in a theory using the loss-of-strength gradient), or complex (the various conditions and their interaction in Deutschian models of integration), or somewhere in between (such as Holsti's study of the issues/stakes of war).

7. One illustration is Most & Starr (1989, chap.6), where the "contending" hypotheses of power preponderance vs. power parity were approached by quickly developing "an experiment that would produce conditions and conclusions regarding the logical properties of the hypotheses," and also help delineate what might be sufficient or necessary conditions.

8. EITM has been specifically directed towards improving "technical-analytical proficiency in Political Science by bridging the divide between formal and empirical analysis" (EITM, 2002, 1; see also the Autumn 2003 Special Issue of *Political Analysis*). The EITM movement has picked up on a number of issues raised by Most & Starr and stressed here. For example: "Specifying a model that links both formal and empirical approaches alerts researchers to outcomes when specific conditions are in place– and is also one of the best ways to determine an identified relationship [emphasis in original] (EITM, 2002, 1-2). Discussing one aspect of EITM, Aldrich & Alt (2003, 310) note: "The result of this mismatch of testing strategic models with methods assuming no strategic interaction is, of course, inaccurate inference." This echoes almost exactly analyses presented in Most & Starr regarding problems studying the causes of war (a minimally dyadic phenomenon) with cross-national data on the characteristics of single states (1989, chap. 3) .