

# Foreign Policy Indicators: Predictors of War in History and in the State of the World Message\*

**J. DAVID SINGER**

*University of Michigan*

**MELVIN SMALL**

*Wayne State University*

---

## ABSTRACT

Despite the promising advances in the construction and use of social indicators, there has been little application to the formulation, monitoring, or evaluation of foreign policy. In the formulation stage—our concern here—predictor or early warning indicators could be very useful. The annual “state of the world” message contains many such predictive indicators of war, but in a purely verbal and intuitive form. Three of these (prior war, relative capabilities, and alliance levels) are converted into operational language and then put to the empirical test. In general, the indicators do not predict war (over the past 150 years) as postulated by the Administration. These tests are, however, very incomplete, and our objective is not to evaluate the Administration’s arguments, but to suggest one way in which indicators could improve the quality of foreign policy formulation.

---

## Introduction

Among the several “movements” of the past decade, one of the more interesting and durable is the social indicators movement. In an effort to cope with the increasing complexities of social planning, as well as with the increasing militance of the socially disadvantaged, practitioners and academics have sought to develop more effective tools for the planning, executing, and evaluating of social policy. Central to that enterprise

---

\* This is a revised and abbreviated version of the paper originally prepared for the 1972 meetings of the American Political Science Association held in Washington, D.C. We want to acknowledge the important assistance of Hugh Wheeler, the comments and help of John Stuckey, Russell Leng, Stuart Bremer, Catherine Kelleher, and Charles Gochman, and the support of the National Science Foundation under grant no. GS-28476X1.

has been the effort to convert many of the vague notions associated with “quality of life” into more precise and operational language. As a result, we now have quantitative indicators which are designed to tap or reflect not only such tangible conditions as “full employment” or “national product,” but more elusive conditions such as the job satisfaction, health care, educational achievement, and environmental quality of a given population.<sup>1</sup>

As these illustrations make clear, the historical origins are found largely in the area of economic activity, with such measurement occurring in the “softer” sectors only in the more recent period. While it is likely, for reasons that will become clear as this paper unfolds, that some of us may have gone overboard in our enthusiasm for these more contemporary indicators, they have already begun to demonstrate their utility.<sup>2</sup> In the foreign policy sector, on the other hand, the utility of indicators has been largely ignored. Outside of the thoughtful exhortations of scholars such as Bobrow, one finds little explicit attention to the kinds of indicators that might be developed, or the ways in which they might improve the efficacy and/or humaneness of national foreign policies.<sup>3</sup> In the article which follows, we hope to focus attention on some of the possible applications of foreign policy indicators and to illustrate the potential that may already lie close at hand.

<sup>1</sup> For some examples and discussions, see: U.N. Department of Economic and Social Affairs, *Report on the World Social Situation*, N.Y.: United Nations (1957–); Raymond A. Bauer (ed), *Social Indicators*, Cambridge: MIT Press (1966); Eleanor Sheldon and Wilbert E. Moore (eds), *Indicators of Social Change*, N.Y.: Russell Sage Foundation (1968); Commission on the Social Sciences, National Sciences Board, *Knowledge into Action: Improving the Nation's Use of the Social Sciences*, Washington, D.C.: National Science Foundation (1969); U.S. Department of Health, Education, and Welfare, *Toward a Social Report*, Washington, D.C.: U.S. Government Printing Office (1969); Otis Dudley Duncan, *Toward Social Reporting: Next Steps*, N.Y.: Russell Sage Foundation (1970); Amitai Etzioni, “Indicators of the Capacities for Societal Guidance,” *The Annals of the American Academy of Political and Social Science*, 388 (March 1970); Fred R. Harris (ed), *Social Science and National Policy*, Chicago: Aldine-Atherton (1970); and Eli Goldston, *The Quantification of Concern: Some Aspects of Social Accounting*, N.Y.: Columbia University Press (1972). One of the earliest efforts to tap a particularly remote social condition is in Norman Bradburn and David Caplovitz, *Reports on Happiness*, Chicago: Aldine (1965). And for one attempt to apply social indicators to the quality of life in the global system, see J. David Singer, “Individual Values, National Interests, and Political Development in the International System,” in Horowitz (ed), *Studies in Comparative International Development*, Beverly Hills, Calif.: Sage (1971). A recent bibliography is Leslie Wilcox et al, *Social Indicators and Societal Monitoring: An Annotated Bibliography*, Amsterdam: Elsevier (1972).

<sup>2</sup> A recurrent theme in the criticism is that expressed by Irving Kristol in “In Search of the Missing Social Indicators,” *Fortune* (Aug. 1, 1969): “These statistics are organized primarily for management purposes. . . . We can account for public money spent [for example] on mental health, but we haven't the faintest idea whether our mental health is getting better or worse. More than that, we don't even have as yet the conceptual apparatus that would enable us to say what we *mean* by mental health, much less permit us to measure it by a series of index numbers.”

Other criticisms—especially of certain economic indicators—are their failure to reflect inequalities in distribution, the assumption that advertising or military expenditures are social “goods,” and their inability to tap the more general concepts of welfare. In response to the latter criticism, one economist retorted: “Producing a summary measure of social welfare is a job for a philosopher-king, and there is no room for a philosopher-king in the federal government.” Arthur Okun, “Should GNP Measure Social Welfare?” *Brookings Bulletin* 8/3 (1971) 4–7.

<sup>3</sup> See Davis Bobrow and Judah Schwartz (eds), *Computers and the Policy Making Community*, Englewood Cliffs, N.J.: Prentice Hall (1968); Davis Bobrow, “International Indicators,” N.Y.: American Political Science Association (Sept. 1969); and Davis Bobrow, “Political and Social Forecasting,” Gaithersburg, Maryland: National Bureau of Standards (March 1970).

## Indicators and their Foreign Policy Applications

In this opening section we cover a number of preliminary points before turning to our specific assignment. First, we attempt to define what we mean by indicators, how they are devised and evaluated, and how they may be differentiated. We then go on to summarize their possible roles in the policy process. Following these preliminaries, we will turn to a more specific application of certain indicators that, while devised for basic research, may turn out to be useful for applied prognosis as well.

### Indicators as Proxies

When we speak of indicators in the context of social phenomena, it is useful to distinguish between two different usages of the term. The first and most familiar usage is in the sense of a proxy, surrogate, symptom, representation, measure, or index. That is, many of the concepts that we use in describing social phenomena do not have easily observed empirical referents. Rather, we make an inferential leap of greater or lesser magnitude from some trace or proxy, which we *do* observe, back to the phenomenon of interest, which often remains unobserved. The condition or event in which we are interested may be unobserved because it has occurred in the past and is gone from sight, or is spread over too broad a space, or has too many component elements, or is sufficiently intangible as to lie beyond the human senses. To illustrate, we may infer the severity of a war from the battle deaths estimates that are compiled afterwards, or the industrial capability of a nation from its steel production, or the material quality of life from infant mortality rates, or the ego strength of an individual from his responses to certain projective tests, or the efficacy of a therapeutic treatment from the patients' average stay in hospital.

In other words, we try to "operationalize" the more elusive concept or variable by devising a procedure through which its presence, strength, or rate of change might be indirectly measured or inferred. If the same operation or procedure, applied to the same phenomenon, gives the same "reading" regardless of who conducts it or when, we say that the index or indicator is a *reliable* one. But devising reliable indicators is only part of the struggle to "observe the unobservable." More difficult and ambiguous is the *validity* of an indicator: the extent to which it really does tap or reflect the phenomenon we claim to be getting at and measuring. Whereas reliability is easily demonstrated by repeated observational tests, validity always remains partially a matter of judgment. There are, however, some ways in which we can go beyond mere assertion as to the validity of an indicator. One is the extent to which an alternative indicator of the allegedly identical phenomenon gives us a set of readings which are highly correlated with those produced by the original indicator. For example, caloric intake might correlate highly (but negatively) with infant mortality, suggesting that both might be valid indicators of material quality of life; or the energy consumption of a nation might correlate highly with its steel production, strengthening our confidence in the validity of both as possible indicators of industrial capability.

A second strategy in the search for validity is to ascertain whether the readings on our indicator conform to what one's model or "theory" would predict. If, for example, our model predicts that the "diplomatic interdependence" of pairs of nations should

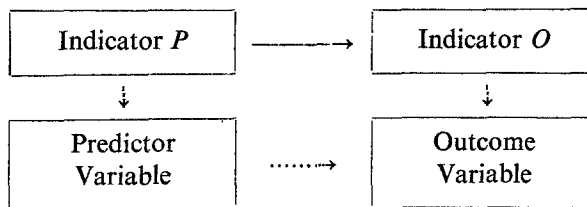
increase as their trade increases, and the number of nations with whom they both have diplomatic relations does indeed increase following a rise in their trade, we have some grounds for believing that the number of nations with whom both members of each pair have relations is a valid indicator of their diplomatic interdependence.

But when all is said and done, the most important test is that of “face validity”: are competent specialists persuaded that the indicator really taps the unobserved phenomenon? And that usually rests, in turn, on the extent to which we believe that: (a) changes in the unobserved phenomenon of interest lead to, or cause, commensurate changes in the value of the indicator; (b) changes in the value of the proxy or trace that serves as the indicator will lead to, or cause, commensurate changes in the actual event or condition of interest; or (c) the values of the variable and its alleged indicator at least rise and fall together, despite the absence of any causal link. In the latter case it may merely be that some third factor exercises an equally strong impact on both the variable and its indicator.

### Indicators as Predictors

Shifting now from the use of indicators as proxies for, or traces of, some less readily observed phenomenon, there is a second and more complex role for social indicators. This is the predictive or early warning role, in which the indicator continues to serve as a proxy or reflection of some unobserved phenomenon, while at the same time it also predicts the changes in the magnitude of some *subsequent* phenomenon. That is, its magnitude is supposed to rise and fall with the magnitude of both the subsequent phenomenon and the *indicator of* that latter event or condition. While an indicator, when used for proxy purposes, is usually measured after (or simultaneous with) changes in the phenomenon it supposedly reflects, it is measured beforehand when used for predictive purposes.

More important than temporal sequence, however, is the extent to which our predictor co-varies with *another indicator* rather than the extent to which it co-varies with some unobserved phenomenon. While this makes it easier to ascertain the degree of co-variation, it also means that we must now deal with the problems of reliability and validity of the indicators at *each* end of the chain. These reliability and validity links are illustrated in the following diagram, in which time runs from left to right and observability runs from top to bottom.



The broken vertical lines leading down from the indicators to the variables emphasize that the latter will often lie below some threshold of direct observability, and their presence or strength will thus have to be inferred from the presence and strength of the indicators which we devise to represent them. The solid line between indicators

*P* and *O* reminds us that it is a simple matter to observe the correlation or co-variation between our predictive and outcome indicators.

But the very simplicity of ascertaining the correlation between indicators can often mislead us into thinking that we have, by that operation, ascertained the predictive, and perhaps even the causal, connection between the unobserved phenomena themselves. If we are merely inferring a *predictive* relationship between the variables, the only (!) threat to that inference lies in the validity and the reliability of our two indicators. But if we want to infer, from a strong correlation between indicators *P* and *O*, that changes in the predictor variable *cause* changes in the outcome variable, all sorts of additional difficulties arise. Since, however, our concern here is not with explanation and causality, but only with the more manageable problem of early warning and prediction, we can side-step these philosophical issues for the nonce. Suffice to say, then, that the weak dotted line conveys the tenuousness of any causal link between the two variables, as well as the extent to which the predictive link between them depends on the quality of their respective indicators.

### **The Policy Uses of Indicators**

Bearing in mind these two meanings of indicator, as well as the dangers of over-interpretation, let us next mention some of the ways in which each type of indicator might be used for policy purposes. They may, of course, be useful not only in: (a) the formulation of policy, but in (b) the monitoring of its execution, and (c) the evaluation of its success.

In the formulation stage, we may use indicators in both their proxy and predictor roles. The former are used to tell us something of the state of the world at the moment, and the direction and rate of change in the conditions and events of interest. The latter may help us to predict which outcomes are most likely to arise out of those conditions, and thus, which of them should be perpetuated (or modified) in order to arrive at (or avoid) some future outcome. In the monitoring stage, we rely primarily on proxy indicators, hoping that they can tell us whether or not the assigned or agreed actions have been executed, and how close certain of the transition conditions are to those expected and thought of as necessary to a successful outcome. And in the evaluation stage, we again turn to proxy indicators to measure the success of the policy, by observing the discrepancy between the ultimate outcome and that which we had predicted and preferred. Moreover, such objective evaluation can be utilized for self-correcting feedback purposes, and provide the basis for change in our predictive models. That is, if the magnitude of the outcome indicator is not as predicted, we know that the appropriate predictor conditions were not those that the model led us to believe, and that other preconditions will have to be established to arrive at the desired outcome. Needless to say, if the indicator is not a valid one, or if the numbers are fictitious, reliance on them can be disastrous. In the Vietnam War, for example, U.S. personnel appear to have not only inflated such indicators as body counts and truck interdictions, but, worse yet, assumed that they were valid indicators of military success.

For predictive indicators to be of much use, however, they must not only satisfy the measurement criteria of reliability and validity. Whereas the historian or political

scientist is often satisfied with the construction of an indicator that merely meets these two criteria—on the assumption that it will in due course be scientifically useful—the policymaker must demand more. He or she must be satisfied that it is also a *dependable* predictor. Either through his own experience, or on the basis of highly credible academic research, he must be persuaded that, in a reasonably large number of cases, the indicator turns out to be a solid predictor. Now, some will ask *how* solid a job of prediction the indicator or index must do before it becomes useful in the policy process. Some officials will insist on 100 percent performance, and will shy away from any indicator unless there is a demonstrated correlation of 1.0 between it and the hoped for, or feared, outcome. The purist would, for example, refuse to take seriously a prediction that increased military pressure by A leads to increased diplomatic resistance from B unless that association has *always* occurred.

This seems overly conservative in two ways. First, it ignores the hard fact that foreign offices are always playing the odds, and no matter how unconscious the operation may be, the estimation of probabilities goes into almost all predictions. Second, this very tendency to think probabilistically reflects an important reality of the political world: very few events or conditions *are* certain. It is not only the state of our knowledge about diplomatic behavior and the changing state of the global system; it is also that *some* degree of randomness will always inhere in these phenomena. Thus, we urge that at this early stage in the development of solid knowledge about international politics, we set more modest and realistic standards. We are not prepared to state what the performance level of an indicator should be in order to take it seriously, but we would point out that, if track records *were* kept, most of us would hope to do better than the two-out-of-three score of U.S. experts on non-routine predictions.<sup>4</sup>

### Predictive Indicators in the State of the World Message

So much, then, for the usefulness of foreign policy indicators in the abstract. Let us shift now to a more specific real world context, in which we find some promising signs of an increased role for such indicators. Reference is to the “state of the world” message inaugurated by President Nixon and his special assistant for national security affairs, Henry Kissinger.

The first of these was presented to the Congress, and the wider world at home and

---

<sup>4</sup> In one of the few efforts to estimate such performance, Jensen found that U.S. State and Defense Department respondents scored 67% and 63% respectively on 25 predictions made in 1965 re: diplomatic events which did or did not occur in the next half decade; see Lloyd Jensen, “Predicting International Events,” *Peace Research Reviews* (1972).

Among the discussions of international event prediction, and tentative efforts to devise predictive indicators, are: Stuart Carter Dodd, “A Barometer of Perceived International Security,” *Public Opinion Quarterly* 9/2 (Summer 1945) 194–200; Gaston Bouthoul, “Les Baromètres Polémologiques,” *Etudes Polémologiques* 1 (1971) 1–26; Rudolph Rummel, “Forecasting International Relations: A Proposed Investigation of Three-Mode Factor Analysis,” *Technological Forecasting* 1 (1969) 197–216; Arthur Vogel, “Toward a Foreign Policy Reporting System,” *World Affairs* 133/2 (Sept. 1970); Uolevi Arosalo, “East-West Trade as a Potential Indicator of International Tension,” Tampere, Finland: *Instant Research on Peace and Violence* 3 (1971) 120–125; Norman Z. Alcock, “The Prediction of War,” Ontario: Canadian Peace Research Institute (1972); and Alan Newcombe and James Wert, *An Inter-Nation Tensiometer for the Prediction of War*, Ontario: Canadian Peace Research Institute (in press). A different approach is the Delphi method; see Norman Dalkey (ed), *Studies in the Quality of Life: Delphi and Decision-Making*, Lexington, Mass.: Lexington (1972).

abroad, at the beginning of Nixon's second year in the White House (18 February 1970). Entitled *U.S. Foreign Policy for the 1970's: A New Strategy for Peace*, it represented both a report on the Administration's first year of foreign policy stewardship and a statement of predictions and preferences regarding the near and middle future. Linked in style and timing to the traditional State of the Union message, it was quickly dubbed—with official encouragement—the “state of the world” message. The second, third, and fourth ones (25 February 1971, 9 February 1972 and 3 May 1973) were similar in form and orientation to the first, and like their predecessor, carried the peace theme in the subtitle.

While there is always some tendency to dismiss such reports as self-serving propaganda and political smokescreen, these state of the world messages obviously have other purposes as well.<sup>5</sup> In addition to assuring the Congress, the media, and the domestic public of the Administration's competence, patriotism, farsightedness, and commitment to peace, the document also meets two other important needs. One is that of communicating a range of signals to other governments, friendly and otherwise, and to some of the world's intergovernmental organizations, in a fairly general and noncommittal fashion. The other is to provide guidance and legitimation to U.S. officials in Washington and in the field.

Finally, although the authors may not have so intended, the state of the world message provides in one place the sort of general overview of world politics that policy analysts may examine for a variety of purposes. One might, by techniques ranging from those of biblical exegesis to quantitative content analysis, search for all sorts of patterns, trends, deviations, inconsistencies, subtle clues, blunt warnings, articulated and unarticulated premises, and so forth.<sup>6</sup> Our purpose is more limited, and our technique quite simple.

Given the persistence of the “peace” theme in all four messages to date, we sought to tap the Administration's collective views as to which particular events and conditions make for peace, conflict, and war in the modern global system and its subsystems. More specifically—and even allowing for the possibility of intentionally misleading statements or deliberately vague interagency compromise phrases—the document offers an excellent opportunity to identify the sorts of indicators used by this Administration, implicitly or explicitly, to predict to, or away from, war. Once these indicators have been “teased out” of the document, we might then be able to convert them into more operational language and then ascertain the extent to which they *have* been dependable predictive indicators in the past. That is, given the tendency of the U.S. and other foreign policy establishments to base their forecasts on “lessons of the past”

---

<sup>5</sup> James Reston, in the *N.Y. Times* (19 Feb. 1970), likened the first one to a maxicoat in that “it is long, it covers a lot of territory, and it conceals the most interesting parts.” Other media reactions included that of David Lawrence in *U.S. News and World Report* (Feb. 28, 1972, p. 92), who saw the third report “as one of the most weighty, most serious, and best argued statements of American foreign policy to be made since the end of the Second World War.” The Palm Beach *Post-Times* (Feb. 12, 1972) on the other hand, described the same one as “an election year exercise in political propaganda,” full of “chest-beating superlatives” set in “gloss and glitter.”

<sup>6</sup> For a systematic, if somewhat superficial, effort to tap the superpowers' operational codes during the late 1950s, via the State Department *Bulletin*, *Pravda*, etc., see J. David Singer, “Soviet and American Foreign Policy Attitudes: Content Analysis of Elite Articulations,” *Conflict Resolution* 8/4 (Dec. 1964) 424–85.

and what "history tells," it is not inappropriate to ask how closely those "lessons" conform to the systematically observed regularities.

In so doing, however, we cannot overemphasize the tentative nature of this comparison. First, despite a conscientious effort to avoid it, we may have misinterpreted the phraseology of these reports, and attributed a position to the Administration which is not justified. This danger is more real when dealing with pre-operational statements, and as foreign policymakers begin to speak in more precise terms, via the explicit use of indicators, the possibility of misinterpretation will, of course, decline. Second, we are looking here at only a small part of the historical evidence: that which has been generated by a single research enterprise, the Correlates of War project at The University of Michigan. When the work of such scholars as North and Rosecrance and their colleagues is brought to bear, a more adequate test of these historical predictions (or post-dictions) will be possible.<sup>7</sup> And, third, given the tentative nature of the results emerging from these projects, it would be premature to claim that we have really pinned down the historical correlations to which we refer.

What we have attempted here, then, is a two-step operation involving these reports. First, we have selected from them a number of predictive statements which embody, if only implicitly, the idea of early warning indicators in the area of war and peace. These are statements which predict that certain preconditions will increase the probability of war for the international system, its regions, or certain specific nations. Second, we have translated these statements into more operational language, using the indicators which had been devised earlier for the Correlates of War project. Rather than spell out the reasoning and procedures behind each of these indicators here, we will cite the book or article in which that information can be found.<sup>8</sup> And, third, we have asked—in a tentative and illustrative fashion only—how regularly the predictor indicators and the war outcomes have been associated in the manner postulated by the Administration. Thus, in each of the following sections, we will offer the predictive statement as it appears in one of the state of the world reports, summarize how we convert the predictor variable into an operational indicator, and examine the extent to which it has been correlated with war over the century and a half from the Congress of Vienna through 1965.<sup>9</sup>

<sup>7</sup> See, as illustrative, Robert North and Nazli Choucri, *Nations in Conflict: Prelude to World War I* (forthcoming), and Richard Rosecrance, *Action and Reaction in World Politics*, Boston: Little, Brown (1963).

Some would say post-diction or retro-diction, but it seems reasonable to speak of pre-diction, even in the past, as long as we have not yet observed and recorded the unfolding of events or the extent to which the predictor and outcome variables did indeed co-vary.

<sup>8</sup> For the indicators, as well as the rationale and procedures, on the incidence of war at the systemic regional, pairwise, and national levels, see J. David Singer and Melvin Small, *The Wages of War, 1816-1965: A Statistical Handbook*, N.Y.: Wiley (1972). As these dates imply, those wars that had not ended, or even begun, by December 31, 1965 are not included: the several Indochina Wars, the Six Day War in the Mideast, the Football War, and those in Yemen, Aden, and Angola.

A useful summary of the overall project will be found in J. David Singer, "The Correlates of War Project: Interim Report and Rationale," *World Politics* 24/2 (Jan. 1972) 243-70.

<sup>9</sup> We must differentiate here between the generalizations which can be drawn from a study of comparable cases out of the past, and the sort of dynamic models which can be generated and tested from an examination of the processes which link these cases together as they unfold across time. As Bobrow reminds us in "International Indicators" op. cit, p. 5, without adequate models "we have no more than descriptive trend plots which lack explanatory power and ignore interaction effects."



## Acceptance of War Leads to More War

Let us turn now to the first of the predictors we have selected for discussion from the state of the world messages. One of the more widely accepted propositions in the folklore of international politics is that war begets war, on the premise that: (a) the victorious initiator of war will seek to repeat his success; (b) that another will be encouraged by the example; or (c) that the defeated party will, alone or with others, move as soon as feasible to settle old scores. In any event, we find in the 1972 report (p. 148) the proposition that “the resort to military solutions, if accepted, would only tempt other nations in delicately poised regions of tension to try the same.”

The predictor variable, like most of those used in diplomacy, can be interpreted (and thus operationalized) in several ways. At one extreme, we can merely ask whether wars come in bunches like epidemics. At the other, we can set up increasingly restricted conditions and ask, for example, whether initiators of wars are more likely to do so again when they have emerged victorious and/or unchallenged; since the number of versions of this hypothesis is quite large, and we hope to examine them in a fairly exhaustive fashion in a subsequent paper, here we will look at just a few versions for purposes of illustration.

In its simplest form, then, we can interpret “resort to military solutions” to mean war in general. Historically, does war lead to more war? During the 150 years from 1816 through 1965, 93 serious international wars began (wars with 1000 or more battle deaths). How often were such wars followed by another within the same or subsequent year?<sup>10</sup> It turns out that 53 (57%) of the 93 wars were so followed by another war. But what does this tell us about war contagion? For instance, how likely is it that those 53 wars would have been followed shortly by another war, on the basis of chance alone?

To test this “null” hypothesis (that it was mere chance), we use the year as our unit of analysis, and ask whether those years during which international war began were more likely to be followed by another war in the same or subsequent year than were those years in which no war began. It turns out that the 93 wars in our study began in 70 different years. Of those 70 war years, 51 saw the onset of one war, 16 saw two wars begin, two years saw three wars, and one year was cursed with the onset of four. Of the 51 single war years, 22 were followed by at least one war in the subsequent year. Those 22, plus the 19 multiple-war years, make a total of 41 years in which war that began in one year was followed by at least one more beginning in the same or subsequent year. That is, while 41 war years were followed by subsequent war years, this also holds for 40 nonwar years. This is a negligible difference and could easily have occurred by chance alone. Turning the question around, we can ask how many

---

<sup>10</sup> Throughout this section, we use this time span of 1–2 years as a reasonable measure of “following soon,” since a longer period would obliterate all distinctions between inter-war intervals. That is, with 93 international wars over the century and a half, we get an average of almost two wars every 3 years, meaning that almost all three year periods would see war following war. And if we eliminate the 43 imperial and colonial wars, looking only at those 50 which were between sovereign states, there would be an average of one war beginning every three years. See *Wages and War*, op cit, Chap. 9.

war and nonwar years were followed by nonwar years. Here it turns out that 29 of the war years and 40 of the nonwar years were not followed by the onset of a new war in the next year. And while this difference is greater than the first—and is in the direction predicted by the President’s message—it is nevertheless far from impressive.

**TABLE 1a**  
**Frequency with which War Years are Followed by War Years, 1816–1965**

		War year?		
		Yes	No	
War follows?	Yes	41	40	$\chi^2 = 1.11$ $Q = 0.17$
	No	29	40	
		—	—	
		70	+ 80	= 150

Using the contingency table (Table 1a) to summarize these figures, we set up two columns, at the bottom of which we show that 70 of the 150 years studied were marked by the onset of one or more serious wars, while 80 were not. Next we set up two rows, so that we can divide the two columns between those years that were followed by war and those that were not. With the four figures (plus the column totals and *their* sum, as a check on our arithmetic, and for later purposes of estimating the statistical significance of our results) we can calculate the extent to which the observed distribution deviates from pure randomness. That is, if the phenomena being examined were completely random, and without some meaningful pattern, all four cells would contain approximately the same number of cases, or about 37 or 38. Thus, by visual inspection alone, we see that the 41–40 difference between war and nonwar years followed by war years is not particularly great, whereas the 29–40 difference seems somewhat sharper. But if we calculate more precisely the strength of association, we find a Yule’s  $Q$  score of only 0.17. And, since we can occasionally get low  $Q$ ’s which are nevertheless statistically significant, as well as high  $Q$ ’s by chance alone, we compute the deviation from randomness as reflected in a coefficient such as  $\chi^2$  (chi-square) and then check it out in a significance table. In this case, the  $\chi^2$  value of 1.11 is so low that we conclude it could easily have emerged by sheer chance.

But this is, after all, much too primitive a test of the Administration’s hypothesis, and a number of more refined ones seem to be called for. To illustrate, given the possibility that there are important differences between the nineteenth and twentieth centuries, it makes sense to test the war-begets-war hypothesis for the two periods separately. In the nineteenth (Table 1b), where 45 of the 84 years (1816 through 1899) saw the onset of war, 31 of them were indeed soon followed by another war, whereas only 14 were not. While not dramatic, this is certainly in the direction postulated by the state of the world message. In the twentieth, on the other hand (Table 1c), where 25 of the 66 years (1900 through 1965) saw the onset of war, only 10 were followed by another war year, whereas 15 were not. This is contrary to the prediction, but again, not overwhelmingly so. Rather than rely on visual inspection of these two sets of figures, though, we once more invoke the  $Q$  and  $\chi^2$  formulae, which take account of the figures in all four cells of each century’s matrix. And, as we would suspect from the very

probable  $\chi^2$  values of 2.00 and 0.26 respectively, the  $Q$  values are also sufficiently low (0.31 and 0.13) to confirm that the historical pattern was quite weak in each of the centuries separately.

**TABLE 1b**  
Nineteenth Century

		War year?		
		Yes	No	
War follows?	Yes	31	21	$\chi^2 = 2.00$ $Q = 0.31$
	No	14	18	
		45	39	= 84

**TABLE 1c**  
Twentieth Century

		War year?		
		Yes	No	
War follows?	Yes	10	19	$\chi^2 = 0.26$ $Q = 0.13$
	No	15	22	
		25	41	= 66

**TABLE 2**  
Frequency with which Wars in which Victim was not aided are Followed by War

		Victim aided?		
		Yes	No	
War follows?	Yes	6	15	$\chi^2 = 0.12$ $Q = 0.11$
	No	7	22	
		13	37	= 50

But this may still be too crude an interpretation of the Administration's proposition. Perhaps the phrase "if accepted" may well be essential to an understanding of the authors' meaning: that the victim of armed attack was not promptly and vigorously joined by others who came to its defense. Thus, we now ask how frequently we find such cases followed by subsequent resort to force. Restricting ourselves to the 50 inter-state wars which mark the 1816-1965 period (thus excluding the colonial and imperial wars, which rarely engaged third parties), our Table 2 matrix shows that 13 were marked by such intervention on the side of those who were attacked, while 37 were not.<sup>11</sup> Since six of the 13 former cases (46%) were soon followed by another war and 15 of the 37 latter were (41%), it would appear to make little difference whether "resort to military solutions" was accepted by the rest of the nations in the system or not. That is, wars in which the victim is left to fight alone are not significantly more likely to be followed by more war than those in which the victim of armed attack *is* joined by others.

<sup>11</sup> Inter-state wars are differentiated from extra-systemic ones in that the former (N = 50) have at least one sovereign state member of the system on each side, whereas the latter (N = 43) see system members fighting against colonies and other less-than-sovereign national entities which do not qualify for system membership.

Alternatively, "accepted" could be interpreted to mean that the initiators were permitted to emerge victorious, and that under *those* conditions, war might beget more war. Here again (Table 3a) the evidence is mixed. Thus, of the 50 inter-state wars in our population of cases, the initiators "won" 34 of them.<sup>12</sup> In 12 of those 34 cases (35%), renewed inter-state warfare broke out in the near future. Conversely, on those 16 occasions in which the initiator did *not* win, the number followed by war was nine (56%). While this is a higher proportional figure, the *Q* value of 0.40 is not quite enough to justify the counter-conclusion that peace is best preserved by permitting the initiator to win.

TABLE 3a

Frequency with which Wars in which Initiator Wins are Followed by War

		Initiator wins?		
		Yes	No	
War follows?	Yes	12	9	$\chi^2 = 1.96$ $Q = 0.40$
	No	22	7	
		34	+ 16	= 50

TABLE 3b

Nineteenth Century

		Initiator wins?		
		Yes	No	
War follows?	Yes	6	6	$\chi^2 = 4.30$ ; $p < 0.05$ $Q = 0.73$
	No	13	2	
		19	+ 8	= 27

In this context, we might again ask whether the historical evidence is more solid in one or the other of our two centuries. In the 19th, which saw 27 of the 50 inter-state wars, 19 were won by the initiator. But only six of those 19, or 32%, were followed by war, while six of the eight, or 75%, which were *lost* by the initiator were so followed. Here, the *Q* coefficient is a rather strong 0.73, and the  $\chi^2$  of 4.30 *is* strong enough to permit a rejection of the null hypothesis;  $p < 0.05$  means that it had a probability of less than 5% of occurring by sheer chance. The interesting thing, however, is that the pattern is quite the opposite of that predicted by the Administration. In other words, for the nineteenth century, wars that the initiator *lost* were more likely to be quickly followed by another.

What about the twentieth century? Here we find that seven of the 15 (47%) wars *won* by the initiator in this century were soon followed by war, while only two of the

<sup>12</sup> We must emphasize that we are only identifying the nations which took the first act of war and initiated military hostilities. This is not always (as in the Franco-Prussian War) the side whose behavior made the war most likely. To ascertain that with any confidence requires a very detailed and reproducible coding of the events leading up to the war, and we are now engaged in that enterprise; see Russell Leng and J. David Singer, "Toward a Multi-Theoretical Typology of International Behavior," (forthcoming). A similar problem exists in identifying the "winner," and for this analysis we merely accept the consensus of the historians who dealt with each of these wars. The Korean War is treated as a "draw"; see *Wages of War* op cit, Chap. 14.

eight (25%) in which the initiator *lost* were so followed—making for a pattern just the reverse of the nineteenth century experience. However, in this century, the relationships, while supporting the Administration's prediction, are not strong enough to permit ready acceptance of the hypothesis that the initiator's being allowed to win is more likely to lead to another war.

TABLE 3c  
Twentieth Century

		Initiator wins?		
		Yes	No	
War follows?	Yes	7	2	$\chi^2 = 1.03$ $Q = 0.45$
	No	8	6	
		15	8	= 23

So far, we have found little support for the notion that wars in which the victim nations are defended by other nations are less likely to be followed by another war than wars in which the victim is *not* defended. The same holds for the notion that defeating the initiator will inhibit the prompt beginning of other wars. But there are further questions that we may ask. For example, among those specific cases in which the victim nation *is* defended by others, are the wars that are won by the initiator more likely to be followed by war? The answer comes as something of a surprise to those of us raised on the ideas of collective defense and collective security. As Table 4 shows, in the 13 cases that were marked by military support of the victim, the initiator was turned back nine times. But in five of those cases, the system saw additional war in the same or following year. On the other hand, if the initiator prevailed despite military support for his victim (it happened four times), there is only one case in which additional war occurred. The  $Q$  coefficient of 0.58 comes close to suggesting that, if peace be the sole concern, it is better for the initiator to win.

TABLE 4  
Frequency with which Wars in which Victim was Aided Successfully are Followed by War

		Victim aided successfully?		
		Yes	No	
War follows?	Yes	5	1	$\chi^2 = 0.31$ $Q = 0.58$
	No	4	3	
		9	4	= 13

Another way to look at this question is to shift from the aggregated systemic level of analysis, and to focus on the specific nations themselves. Here we ask whether those nations which initiated war turn out to initiate another war within the following ten years. As Table 5 indicates, it matters a great deal whether the initiator emerged victorious or not. That is, of the 73 nations that were on the initiating side, only 13 of them were initiators again within the decade. But 12 of those 13 had won the first war, and the high  $Q$  and  $\chi^2$  values show that this is a very strong and statistically

significant relationship; that is, such a  $\chi^2$  had less than a 2% probability of occurring by chance alone.

Further evidence as to the effect of victory is seen in Table 6, where the coefficients are even more persuasive. Here, we see that of the 98 nations that were on the side which was attacked, only 11 of them turned around and initiated war within the decade. But again, 10 of those 11 future initiators, even though often caught unprepared, nevertheless emerged on the winning side.

**TABLE 5**  
**Frequency with which Initiators of War Soon Initiate Another War**

		Initiators win?		
		Yes	No	
Initiates	Yes	12	1	$\chi^2 = 5.82; p < 0.02$ $Q = 0.80$
another war?	No	34	26	
		46	27	= 73

**TABLE 6**  
**Frequency with which Victims of Attack Soon Initiate Another War**

		Victim wins?		
		Yes	No	
Initiates	Yes	10	1	$\chi^2 = 6.77; p < 0.01$ $Q = 0.82$
another war?	No	43	44	
		53	45	= 98

In sum, these two analyses suggest that it is not so much that war begets war, since very few nations initiate wars in the decades following prior war experience, but that *victorious* war begets war, and this finding is certainly consonant with the Administration's hypothesis. More generally, we see that the simple proposition is in fact a rather complex one, and that the alternative ways of interpreting and testing it (via many alternative indicators) lead to rather different conclusions. The evidence, then, is far from complete, and in a subsequent paper we intend to examine the epidemiology question in considerably greater detail.

### **Weakness Leads to War**

A close corollary of the proposition that the global community's *acquiescence* in the initiation of war will lead to more war is that which tells us that relative *weakness* leads to war. This may be interpreted as a national level prediction, with weaker nations being more vulnerable and therefore more likely to *have* to fight in self defense; it may also be interpreted in systemic terms, with weakness on the part of the "peace loving" (or "status quo") nations serving as a temptation to the stronger. This type of predictive statement finds repeated expression in all four of the Administration's reports on the state of the world. It appears in particularly crisp language in the first of these: "Peace requires strength. So long as there are those who would threaten our vital interests and those of our allies with military force, we must be strong.

American weakness could tempt would-be aggressors to make dangerous miscalculations” (p. 4). And the theme is echoed in 1972: “American weakness would make no contribution to peace. On the contrary, it would undermine prospects for peace” (p. 155).

In discussing the relationship, past or predicted, between war-proneness and national power or strength, we quickly come face to face with the familiar question of “what is power?” In a forthcoming volume, we will discuss that question in painstaking detail, but here we must be brief.<sup>13</sup> First, there is the difference between a nation’s power *base* or *potential* on the one hand, and military preparedness or other indices of force *in being*, on the other. Second, the bases of power have changed across time. Third, a nation’s capacity to exercise influence may range considerably across space. The indicator we employ here is a very general one, and that limitation should be kept in mind as the discussion unfolds.

We use a straightforward measure in which we combine six factors: steel production, energy consumption, urban population, the square root of total population, armed forces size, and military expenditure. For more thorough analyses, we use several different combinations and weightings of these six, but here we will merely compute each nation’s share of each element’s distribution in the system, and then compute the average of its percentage shares. Thus, if a nation has 8% of the system’s armed forces, steel production, and urban population, and 12% if its military expenditures, energy consumption, and total population (converted to its square root), its composite score would be 10%. In the analysis at hand, restricted only to the major powers for the 150-year period, the rankings and scores turn out to be remarkably close to those that a diplomatic historian might expect; in other words, the face validity of the indicator—as a reflection of general power potential—appears to be quite high. (For a more restricted test, and perhaps one more appropriate to the Administration’s argument, we might want to use the military dimensions only, or better yet, an indicator which reflects the fraction of overall capabilities that has been allocated to military preparedness.)

Using this composite indicator, *does* it turn out that major powers are more likely to get into war, or to be attacked, when they are: (a) nearer the bottom of the capability scale, or (b) experiencing a decline in relative power? In Table 7, we present a rough summary of where each major power stood vis-à-vis the others in that subset on the eve of its serious inter-state wars, as well as the direction in which its capabilities tended to be moving, in regard to the average score of the majors at that time.

In summarizing so large a body of information in such a table, we must emphasize several limitations. First, as already noted, our composite capability measure is still a tentative one whose validity is far from conclusive. Second, we are still refining the data on which the indicator is based. Third, specification of initiators and defenders in war is by no means self-evident, and as we complete our historical reconstruction of the events leading up to each war, there may be revisions as well as refinements in our classifications. With these caveats in mind, we look at Table 7, which lists the

---

<sup>13</sup> The data acquisition procedures, sources, theoretical arguments, and resulting scores will be found in J. David Singer *et al*, *The Strength of Nations: Comparative Capabilities since Waterloo*, (forthcoming).

capability status of each major power on the eve of its entries into inter-state wars over the 150 years under study.

Of the total of 59 war entries, 34 occurred when the combatant state was weaker than the average major power, while only 25 occurred when it was above the average on the composite capabilities indicator (see column totals in Tables 7 and 8). To the extent that we have a valid indicator of strength and that the average score constitutes a reasonable cutting point between the weaker and the stronger, the Administration's prediction finds some historical support. Next, we ask whether that pattern is reinforced by the association between the nation's war experiences and the rise, stability, or decline in their capability scores (see totals in table 7 and rows of table 8). Here we find virtually no difference: 30 occur when the major power's score is falling, and 29 when it is either rising or holding steady. As to the combined "effect" of both scores, the 13-12 and 16-18 figures reinforce the impression that present strength is moderately important and that direction of change is of minor consequence. On the other hand, the U.S. experience is unique in that all three of its wars (since joining the major power class) began while it was in the upper half, but on the decline.

It is, of course, one thing to participate in a serious war and quite another to be more or less *compelled* to do so, as a result of another power's initiation of hostilities. Thus, we now move on to refine our inquiry, and ask how frequently the *victims* of military assault by another major power were on the high or low side of the major

TABLE 7  
Relative Capabilities of Major Powers Prior to Their Entry into Inter-state Wars

	Above major power average		Below major power average	
	Steady or rising	Falling	Steady or rising	Falling
France	Franco-Span, 1823 Navarino Bay, 1827 Crimean, 1854 Franco-Mex, 1862 Franco-Pruss, 1870	Roman Repub, 1849 <i>Ital Unif</i> , 1859 Sino-French, 1884	<i>WW I</i> , 1914	<i>WW II</i> , 1939 Korean, 1951 Sinai, 1956
Britain	Anglo-Persian, 1856	Navarino Bay, 1827 Crimean, 1854 <i>WW I</i> , 1914		<i>WW II</i> , 1939 Korean, 1950 Sinai, 1956
Germany	<i>WW I</i> , 1914 <i>WW II</i> , 1939		2nd Schles-Holstein, 1864 Seven Weeks, 1866 <i>Franco-Pruss</i> , 1870	1st Schles-Holstein, 1848



TABLE 7—continued

Aust-Hung			Ital Unif, 1859	Austro-Sardin, 1848 Roman Repub, 1849 2nd Schles- Holstein, 1864 <i>Seven Weeks, 1866</i> <i>WW I, 1914</i>
Italy			Italo-Roman, 1860* Italo-Sicil, 1860* <i>WW I, 1915</i>	Seven Weeks, 1866 Italo-Turk, 1911 Italo-Ethiop, 1935 WW II, 1940
Russia	Crimean, 1853 Russo-Jap, 1904 <i>Russo-Jap, 1939</i> Russo-Finn, 1939 <i>WW II, 1941</i>	Navarino Bay, 1827 Russo-Turk, 1828 Russo-Hung, 1956	<i>WW I, 1914</i>	Russo-Turk, 1877
U.S.		<i>WW I, 1914</i> <i>WW II, 1939</i> Korean, 1950		
China			Korean, 1950 Sino-Indian, 1962	
Japan			Russo-Jap, 1904 <i>WW I, 1914</i> Manchurian, 1931 Russo-Jap, 1939 WW II, 1941	Sino-Jap, 1937
Totals	13	12	16	18

- Notes: (1) Italics—Attacked by another major.  
 (2) Dates—Year of entry.  
 (3) \*—New in system and no prior measurement of its capability.

power average, and whether their shares of the capability pie were rising, holding steady, or falling. As the  $Q$  value of 0.20 and the  $\chi^2$  of 0.09 in Table 9 show, this version of the Administration's proposition stands up even less well. Among the nine major power victims, five were above average in capability while four were below, and five were steady or rising while four were falling. These cell entries are as near equality as possible, given the  $N$  of 9. Historically, therefore, we conclude that when major powers have fought one another, the victim nation is no more likely to be below the average major power capability scores than above, and no more likely to be falling in its share of the composite capabilities than it is to be rising or holding steady.

**TABLE 8**  
**Relative Capability Frequencies of Major Power War Participants**

	Above	Below	
Steady or rising	13	16	$\chi^2 = 0.10$ $Q = 0.14$
Falling	12	18	
	—	—	
	25	+ 34	= 59

**TABLE 9**  
**Relative Capability Frequencies of Major Power Victims of War Initiation**

	Above	Below	
Steady or rising	Crimean Russo-Jap (04) Russo-Jap (39)	WW I Fr-Pruss.	$\chi^2 = 0.09$
Falling	Ital. Unif. Korean	WW II Seven Weeks	$Q = 0.20$
	5	+ 4	= 9

Once again, though, there are alternative ways to operationalize, and put to the test, these state of the world propositions. In this case, it may be more important to compare those who were attacked to their *attackers*, rather than to *all* the other major powers; and the figures in Table 10 show that such a modification does make a modest difference. That is, in the nine major-major wars, we find that the victim was weaker than the attacker in just over half the cases, and that is a negligible difference. On the other hand, all of the four that *were* stronger than their attackers were *declining* in relative capabilities, as were three of the five that were weaker, and this pattern lends some credence to the state of the world argument. To put it in its strongest sense, we find no cases of a major power being attacked by another when the first was both stronger than, and rising vis-à-vis, the potential attacker.

**TABLE 10**  
**Relative Capability Frequencies of Victims  
vs. Initiators in Major Power Wars**

	Victim stronger	Victim weaker	
Steady or rising		Crimean Fr-Pruss.	$\chi^2 = 0.51$
Falling	Ital. Unif. Russo-Jap (04) Russo-Jap (39) Korean	Seven Weeks WW II WW I	$Q$ cannot be calculated when any cell is empty

This does not, of course, wrap up the argument. So far, we have only examined the capability-war relationship when war *did* occur. It is equally crucial to examine those cases in which war did *not* occur. Thus, in Table 11, we summarize all the half-decade experiences of all the major powers, and ask two simple questions.

TABLE 11

Frequencies with Which Major Powers were Below Average in Capabilities and Entered into War

		Below average?		
		Yes	No	
Entered war?	Yes	31	21	$\chi^2 = 1.37$ $Q = 0.19$
	No	63	63	
		—	—	
		94	+ 84	= 178

First, how many major power half decades are characterized by an above average military-industrial capability, and second, how many of them were followed by entrance into one or more wars of any type? As the frequency distributions show, there were somewhat more below average than above average cases (94 to 84), but many more peaceful than warlike cases (126 to 52). More important, when we put the two sets together, we find that capability makes very little difference. Whereas more of the warlike cases (31 to 21) involve below average majors, the nonwar cases include just as many above as below capability scores. And as the weak  $Q$  value of 0.19 and the  $\chi^2$  of 1.37 make clear, the pattern is much too close to random to justify any interpretation either way.

Similarly, if we look (Table 12) not at capability scores for these major power periods, but ask whether or not they were *preceded* by a fall in strength, the same essential picture emerges. That is, of the 80 nation periods preceded by a fall in capability, 29 (33%) were followed by war involving those majors which fell in relative capability, while 20 (25%) of those 80 that rose in capability were also followed by war. Again, the  $Q$  value of 0.18 reveals only very weak support of the proposition, although in the predicted direction.

TABLE 12

Frequency of War Entry by Major Powers at Times of Declining Relative Capability

		Capability falling?		
		Yes	No	
Entered war?	Yes	29	20	$\chi^2 = 1.18$ $Q = 0.18$
	No	60	60	
		—	—	
		89	+ 80	= 169

In sum, we have found very little support for the Administration's "weakness leads to war" proposition. While our limited historical experiment tends to confirm the belief that when major power wars occur, the victim nation was falling in capability relative to the initiator, we cannot say much about the likelihood of *getting into* war from knowing anything about capability and direction of change. That is, major powers seem to be as likely to get into war when their capability is above the major power average as when it is below it, and they seem as likely to get into war when their capability is rising as when it is falling.

## Alliances Help to Deter War

As intimated in the previous sections, the likelihood of war and the nature of its outcome will depend not only upon the relative strength of the would-be initiator and its victim. The capabilities of allies can often be conclusive. But it is usually too late to begin forming alliances once war has begun, or even on its eve; as a matter of fact, of the 177 formal alliances established during the 150 years under study, only eight were consummated during, or within the three months preceding, hostilities. This tendency, by itself, suggests that practitioners have generally shared the Administration's view that "we must build an alliance strong enough to deter those who might threaten war" (1970, p. 27).

Such an inquiry requires three different foci. First, is the international system less war-prone when alliance levels are high? Second, are the nations which join into most alliances also the ones which experience the least war? Third, does high alliance involvement on the eve of war make a nation less likely to be drawn into that war?

Before examining the historical correlations between alliance levels and war, we should summarize the operations and reasoning behind our indicators. Despite the relative ease of identifying all formal alliances (at least after the archives are opened) and differentiating among their types, ours represents the first published effort to convert that diplomatic information into machine-readable, quantitative, indicators.<sup>14</sup> First, we only include *written* treaties of alliance between and among sovereign states. Second, we differentiate among: (a) defense pacts, in which the signatories contract to *fight alongside* one another if either is "attacked"; (b) neutrality or nonaggression pacts, in which the obligation is to *not fight against* the other; and (c) ententes, which oblige the signatories to *consult* in the event, or imminent likelihood of, hostilities. Third, we exclude: (a) such highly asymmetric alliances as in treaties of guarantee; (b) collective security arrangements such as the League and the U.N., in which the potential attacker may well be one of the signatories; and (c) general declarations of nonviolent behavior, such as the Kellogg-Briand Pact or the Geneva Conventions.

With all alliances, members, and dates recorded, we then combine the information in order to measure, for each year, what percentage of the *system's* members are in one or more alliances of each type; these are our alliance *aggregation* indicators. For each *nation*, we ascertain the number of alliance commitments of each type it has with different classes of nations each year, in order to get its alliance *commitment* index.

Looking at the question first in its most general form (before going on to a more refined analysis), we ask whether high alliance aggregation levels do indeed lead to low levels of war in the system, and vice versa. In Table 13 we show the "product-moment" correlations between the various alliance and war indicators for the entire time span, and then for each of the centuries separately. If most of the high alliance aggregation periods were followed by high amounts of war, and the low alliance periods were followed by low amounts of war, these correlations would be positive, and would range

<sup>14</sup> For details of our coding procedures and the resulting data sets, see J. David Singer and Melvin Small, "Formal Alliances, 1815-1939: A Quantitative Description," *Journal of Peace Research*, 3/1 (Jan. 1966), 1-32 and Melvin Small and J. David Singer, "Formal Alliances, 1816-1965: An Extension of the Basic Data," *Journal of Peace Research*, 3 (1969) 257-282.

from about +0.50 up to a maximum possible value of +1.00. Conversely, if these high alliance aggregation levels were followed by the *least* warlike three-year periods, the correlation coefficients would show a *negative* sign, and again could in principle approach a value of -1.00. And if these alliance and war levels are temporally associated in a largely random fashion, the coefficient values would be close to zero. Further, as with the contingency table analyses, we are interested in both the strength of the association and the likelihood—reflected in the “statistical significance” level of the correlation coefficients—that the higher ones did not occur by chance alone.

TABLE 13  
Correlations Between Alliance Aggregation and Onset of War

Predictors	No. of wars	Nation months	Battle deaths
	%	%	%
<i>Total Time Span</i>			
nations in any alliance	0.00	0.27	0.26
nations in defense pacts	-0.01	0.05	-0.01
majors in any alliance	0.01	0.20	0.20
majors in defense pacts	0.00	0.11	0.09
<i>Nineteenth Century</i>			
nations in any alliance	0.03	-0.09	-0.24
nations in defense pacts	0.00	0.10	-0.23
majors in any alliance	-0.21	-0.18	-0.34
majors in defense pacts	-0.16	-0.28	-0.45*
<i>Twentieth Century</i>			
nations in any alliance	0.00	0.23	0.21
nations in defense pacts	-0.02	0.05	0.04
majors in any alliance	0.34	0.31	0.28
majors in defense pacts	0.19	0.12	0.04

\* Less than 5% probability of occurring by chance alone.

Turning to Table 13, then, we find that the patterns are very close to random for the *total* time span; regardless of which indicators we use for the predictor and outcome variables, the highest correlation we find is a weakly positive one of 0.27. Skipping to the twentieth century alone, the signs are again by and large positive, and thus in the direction opposite to that predicted by the Administration, but with a somewhat more consistent pattern. For example, when most of the *major* powers are in one or more alliances, the following periods tend to be fairly warlike, as measured by the frequency of wars begun as well as by their magnitude in nation months or their severity in battle-connected deaths. But even the 0.34 correlation had more than a 5% probability of occurring by chance alone.

Reverting to the nineteenth century, however, we find a reasonably clear congruence between the state of the world report's propositions and the historical record. Here, most of the correlation coefficients are negative, but only one of the 12 had less than a 5% probability of emerging by chance alone. So far, then, the Administration's view

is closer to that found in the previous century than in the current one, but none of the patterns is particularly strong.

Turning now from a consideration of the alliance-war relationship in the entire inter-state system, we examine the more restricted grouping of major powers only. In this case, the query is whether those majors which had the highest levels of alliance participation during their entire tenure in the system were also those which experienced the lowest incidence of war involvement. Here, we can say nothing about a possible causal connection because these nations could, for example, have had most of their alliance participation in one time period and their war involvement quite a few years later or earlier. In any event, if the administration's hypothesis is correct, most of the correlation coefficients should be strong and negative. As Table 14 indicates, the alliance-war relationship is again quite confused.

**TABLE 14**  
**Correlations Between Major Power Alliance Involvement**  
**and War Experience**

Predictors	No. of wars per year	Months of war per year	Battle deaths per year
All alliances	0.60	-0.39	0.13
All Defense pacts	0.42	-0.45	0.07
All Neut. pacts	0.25	-0.07	-0.47
All Ententes	0.67*	0.01	0.49
Alliances with majors	0.00	-0.26	-0.49
Defense with majors	-0.03	-0.37	-0.28
Neut. with majors	0.11	-0.23	-0.58
Ententes with majors	0.00	0.08	-0.66
Alliances with minors	0.68*	-0.35	0.31
Defense with minors	0.51	-0.43	0.16
Neut. with minors	0.28	0.00	-0.40
Ententes with minors	0.60	-0.01	0.62

\* Less than 5% probability of occurring by chance alone.

On the other hand, there is a very consistent and rather strong *positive* association between the *number* of alliances each of the majors had and the number of wars that each experienced, controlling for the length of time that they belonged to the major power subsystem. But this seems to hold only for alliances with *minor* nations, since these indices of alliances with other *majors*—and these were quite rare—show an essentially random association with their war proneness. Further, it matters considerably whether our indicator is the *frequency* with which allied powers get into war, or the *amount* of war they experience. That is, when we correlate their alliance involvement with the magnitude and severity of war which they experience, we find most of the coefficients to be negative, as postulated. However, none of these is sufficiently strong to have had a less than 5% probability of occurring by chance alone.

Despite the consistency of these positive and negative signs, we must avoid over-interpreting them. As noted earlier, such coefficients tell us nothing of the temporal sequence; once more, a nation could have most of its alliance involvement in one historical epoch and its wars in another. Thus, a more realistic test for the proposition that alliances deter war is to examine the extent to which high alliance involvement is closely *followed* by low war experience for the nations, and vice versa.

What we examine next is the frequency with which major power membership in one or more alliances in a given year is followed by entry into inter-state war during that same year or within the two following years. As the matrix in Table 15 shows, there is virtually no discernible pattern. Of the 648 major power alliance membership years, 116 (18%) are followed by war entry of the alliance members, while 532 (82%) are not. Even though this looks as if alliance membership reduces war experience, we next note that of the 239 nation-years *not* marked by alliance membership, 44 are nevertheless followed by entry into war and 195 are not. These percentages are also 18% and 82% respectively, identical to those for years in which these nations were in alliances. The similarity of these ratios is confirmed when we compute the  $\chi^2$  and  $Q$  coefficients, each of which is negligible. In other words, membership in alliances seems to have had little effect on the historical likelihood of major powers getting into war or remaining at peace.

**TABLE 15**  
Frequencies with which Majors Belonged to Alliances  
and then Entered into War.

		In alliance?		
		Yes	No	
Entered war?	Yes	116	44	$\chi^2 = 0.03$ $Q = 0.02$
	No	532	195	
		648	+ 239 = 887	

**TABLE 16**  
Frequencies with which Majors Joined Alliances  
and then Entered into War

		Joined alliance?		
		Yes	No	
Entered war?	Yes	31	129	$\chi^2 = 2.6$ $Q = 0.18$
	No	104	623	
		135	+ 752 = 887	

Finally, the historical association between alliance membership and war participation may be examined in terms of the effect of *joining* an alliance rather than merely *being* in one. In other words, does the *formation* of alliances turn out to predict decreases in the war proneness of nations, particularly the major powers? In Table 16, we present the familiar four sets of frequencies, beginning with the number of years in which any of the majors did or did not join a new alliance. We then divide those frequencies according to whether those joining and nonjoining years were or were not followed by entry into an inter-state war. We reiterate in passing that we do not count those few alliances which were entered into during, or less than three months before,

a war; this is to eliminate the confounding effects of those alliances which are created in order to fight rather than deter, a war.

As the  $Q$  and  $\chi^2$  coefficients indicate, the pattern is far from dramatic, but what association we do find is in the direction opposite to that predicted. Of the 135 alliance entries by major powers, 31 (23%) were followed by their entry into war, and 104 (77%) were not. However, note that of the 752 nation-years which did *not* see major power alliance entries, only 129 (17%) were followed by entry into war by those nations, while 623 (83%) were not. In other words, those powers which entered into alliances were somewhat more likely to subsequently get involved in war than those major powers which did not. But the association is far from a statistically significant one.

We have, therefore, some tentative evidence that, historically, alliances have not prevented war. As a matter of fact, our more realistic tests of the temporal sequence between alliance involvement and war suggest that, in the case of alliance *membership*, major powers are as likely to get into war while participating as not, and, in the case of alliance *formation*, major powers are more likely to get into war subsequent to joining alliances than not. At best, then, it seems that the creation of alliances has done little to keep major powers out of war.

## Conclusion

In our effort to illustrate ways in which contingent predictions (“if —, then, —”) in foreign policy might be improved via the use of operational indicators, we have merely scratched the surface. First of all, we have selected such predictions from only one set of documents reflecting a single nation and a single government. Second, we have only selected a fraction of those which could be examined in the light of a small part of a single research project. Third, all of the contingent predictions are of a simple bivariate nature, with a single predictor variable; and while many practitioners and researchers tend to think in essentially bivariate terms, a moment’s pause reminds us that few outcomes in international politics are likely to be “determined” by a single factor.

Furthermore, we have addressed ourselves primarily to the *prediction* problem in policymaking and have ignored many other critical issues, ranging from the decisional setting through effective policy implementation. Nor have we wrestled with the differentiation between factors which might have a great deal of predictive or explanatory power and those which, while not theoretically powerful, may offer greater opportunity for conscious and timely human intervention.<sup>15</sup>

On the other hand, this modest exercise should convey an idea of the possibility that we may indeed “learn from history.” To begin, if we can frame our policy prediction statements in more precise language, we are already part way to the construction of reliable and valid indicators. If we can begin to approach history in a systematic and rigorous fashion, rather than merely ransack it for arguments and analogies that are

---

<sup>15</sup> Neither have we examined the policy recommendations that might flow from these tentative findings. That responsibility we will take up in future papers, when focusing more thoroughly on one set of problems at a time.



convenient at the moment, we can begin to accumulate a fair amount of existential and correlational knowledge. If we can assimilate that knowledge through a multi-theoretical taxonomy, we can begin to develop modest islands of explanatory knowledge. If that knowledge can be integrated into alternative theories, these can be put to further historical test. And, from the more successful and accurate theories, additional predictions can be made. A sequential scenario such as this might, in due course, lead to the conditions sought by the conscientious policymaker. These were put nicely in the first state of the world message, "Our actions must be the products of thorough analysis, forward planning, and deliberate decision. . . . We must know the facts: intelligent discussions . . . and wise decisions require the most reliable information available. . . . We must know the alternatives: we must know what our real options are, and not simply what compromise has found bureaucratic acceptance" (pp. 17-18).<sup>16</sup>

Having sounded this moderately sanguine note, however, it now behoves us to back off and face up to some of the more critical issues that confront the applied scientist. First, there is the question of whether social scientists *should* be concerned with the policy implications of their research. Should we actively emphasize those implications, ignore them, or intentionally conceal them, to take the three dominant views? Further, should we make them available to some governments and not to others? To some political parties and not to others? To some agencies or bureaus and not to others? In the short run, of course, these remain fairly academic questions, given the paucity of our correlational and predictive (not to mention explanatory) knowledge, and the skepticism of foreign office types the world over.<sup>17</sup>

But as (and if) the proposed trend continues, and our knowledge base becomes more solid, these can emerge as increasingly salient questions. To respond in an indirect fashion, we have no illusions that all the disasters that befall mankind are a consequence of ignorance or incompetence; nor do we urge that if we merely understood more fully the consequences of our actions, life would be less nasty, brutish, or short. Similarly, we do not suggest that there is little difference between the "good guys" and the "bad guys" in classifying nations, regimes, agencies, or individuals. But we would nevertheless argue that an applied social science is an essential concomitant of any related efforts to move toward a world of peace, prosperity, and justice.

Further, the same extension and diffusion of knowledge that can permit *policy-makers* and their interpreters to become more sophisticated and insightful regarding the issues which confront them, can also serve to disenfranchise the *general citizen*. Whether cheer-leader, nay-sayer, or merely the victim of governmental decisions, the

---

<sup>16</sup> In the first issue of the *Journal of Conflict Resolution* (1957, p. 94), Quincy Wright proposed the establishment of a "world intelligence center" among whose missions would be the measurement of "the changing atmosphere of world opinion [and] the changing condition of world politics." Such a use of foreign policy indicators, whose scores would be published on a weekly or monthly basis by an independent global institute, could appreciably enhance the accuracy and credibility of the facts and help clarify the range of alternatives to which the Administration alludes.

<sup>17</sup> As to the problems of communication between social scientists and the policymaker, see Philip Burgess, "International Relations Theory: Prospect, 1970-1995," Los Angeles, Calif.: American Political Science Association (1970) 9-14; and Charles Hermann, "The Knowledge Gap: The Exchange of Information between the Academic and Foreign Policy Communities," Chicago: American Political Science Association (1971).

man in the street can become a more critical and knowledgeable participant in the governmental process. And, in doing so, he will increasingly blunt one of the major weapons of demagogues, bigots, and warmongers: the combined ignorance which permits the leader, as well as the led, to believe the most ill-founded and foolish propaganda.

This consideration leads us, with some reluctance, to the unattractive issue of pseudoscience or "scientism." Not only the general citizen and the politician, but the specialists in foreign and defense ministries are always in danger of being taken in by research that *looks* like science. To some extent, the universities, but more often the "think tanks" and industrial research firms are well populated with "number jugglers" whose reports have a decidedly scientific aura to them. The background factors that make for shoddy work are all too familiar: (a) today's strategic analyst or Middle East specialist was yesterday's expert on traffic safety or water pollution; (b) a large fraction of the organization's business comes from a mission-oriented client with the consequent political pressures; and (c) time pressures often guarantee a quick and dirty job on contract research.<sup>18</sup>

But most often, the source is simple ignorance at the client's end as well as at the researcher's end. This need not be surprising, given the durability of the "three-culture problem" in today's world. Sir Charles Snow reminded us of the gulf between the scientific culture and the humanistic culture, but he neglected to emphasize that those of the former class are no less ignorant of the social sciences than those of the latter; engineers, physicists, and systems analysts often have even more erroneous notions about social systems (and how to study them) than poets, lawyers, or businessmen.

Thus, we conclude with an emphasis on not only the familiar need for more research, but the less familiar one of the need for a very different kind of "education for world affairs." In the primary and secondary schools, in colleges and graduate schools, and perhaps more critically, in the public and private discussion of public policy questions, it is essential that we move out of the prescientific era. Superstition was of little use in understanding the weather or estimating the strength of a tunnel, and medieval notions of human physiology were of little help in diagnosing and curing disease. Why should we hope that such primitivism will suffice in the solution of foreign policy problems? Given the destructive effects of bureaucratic politics, as well as the mix of good intentions and bad, there is no need to compound the felony with so large a measure of ignorance. Surely we can do better.

---

<sup>18</sup> These are, of course, impressionistic judgments that could and should be put to the test; a more affirmative view is in Roger Levien, "Independent Public Policy Analysis Organizations: A Major Social Invention," Santa Monica, Calif.: Rand Corporation (Nov. 1969).