CAUSAL INFERENCES AND THE USE OF FORCE: A CRITIQUE OF FORCE WIT--ETC(U)

MAY 80 S M WALT

UNCLASSIFIED CNA-PP-279

END

DATE 7 80

DTIC
CAUSAL INFERENCES AND THE USE OF FORCE: A CRITIQUE OF FORCE WITHOUT WAR

Stephen M. Walt
The ideas expressed in this paper are those of the author. The paper does not necessarily represent the views of either the Center for Naval Analyses or the Department of Defense.
CAUSAL INFERENCE AND THE USE OF FORCE:
A CRITIQUE OF FORCE WITHOUT WAR

Stephen M. Walt

Institute of Naval Studies
CENTER FOR NAVAL ANALYSES
2000 North Beauregard Street, Alexandria, Virginia 22311
INTRODUCTION

The use of the armed forces as a political instrument in situations short of war has attracted considerable attention in recent years. Because major wars have been infrequent and are likely to remain so, the armed forces of the leading states are likely to be most frequently used in discrete demonstrations intended to have political effects. Because the issues surrounding the political use of military force are important -- both in terms of procurement decisions and the potential employment of such forces in international disputes -- the methods used to address this topic must be carefully chosen. Failure to do so can lead to inconclusive findings and potentially misleading policy recommendations, while appropriate methods and careful analyses can enhance our understanding of international politics and contribute to effective policy choices.

A recent Brookings Institution study, Force Without War, illustrates both the difficulties involved in assessing the role of the armed forces in purely political situations, and the problems that result from inadequate attention to the methodological requirements one's substantive interest imposes. The study has attracted a great deal of favorable attention since its publication, despite the
fact that it contains a number of major methodological flaws. The apparently widespread failure to recognize these flaws suggests a basic misunderstanding of the requirements for sound analytical work in international relations.

This essay attempts to alleviate this problem in two ways. First, I shall briefly outline the basic requirements for evaluating the importance of the armed forces as a purely political instrument. To do so requires a brief recapitulation of the fundamental logic of scientific inquiry, and this should already be familiar to many students of behavioral science. It is necessary because the task of evaluating the importance of military force is as much a theoretical problem as it is an empirical one. As a result, a clear understanding of how causal inferences are made is required. Second, I shall discuss at some length the methods used in _Force Without War_, and shall attempt to reveal the major shortcomings in both the methods used and the conclusions reached. Readers who are already familiar with the methodological problems of causal inference and theory construction may wish to move immediately to the section beginning on page 14.
As Thomas Schelling said in the preface to *Arms and Influence*:

One of the lamentable principles of human productivity is that it is easier to destroy than to create. A house that takes several man-years to build can be burned in an hour by any young delinquent who has the price of a box of matches.

In what follows here, readers may find too little attention given to the undeniably major intellectual effort which informs almost every page of *Force Without War*. The prose is lucid, the attempt to grapple with complex issues sincere, and the use of a multi-level approach to data analysis (aggregate methods as well as case studies) is refreshing. The critique offered in this essay is, ultimately, an attempt to help those of us who struggle with questions of foreign policy and instruments of international influence push back the areas of uncertainty in this business a little further.

CAUSATION, THEORY, AND THE USE OF FORCE

Laws and Theories

Whether in physics, economics, or international politics, understanding a given phenomenon means that one is able to identify the causal forces that work to produce different outcomes. The central scientific question is always "why does this occur?", and the merits of any scientific in-
quiry are judged by the ability of the analyst to demonstrate the causal relationships that determine the observed events.

Understanding the role of military force as a political instrument is subject to these same requirements. In order to know how to best examine this issue, we need to know how to determine causality. What are the requirements for making credible causal inferences?

The first requirement is simplification. No event can be completely described, so the analyst is forced to select those traits believed to be important, while ignoring others. This task is accomplished by selecting information according to general concepts that apply across the range of events we are trying to explain. Concepts are simply classifications into which items may be grouped according to certain shared characteristics. In this way, aspects of different events can be viewed as representatives of a more general family, and possible causal relations within the different events can be explored.

Concepts must be precisely defined. After all, if we are to say that "A causes B," we must be clear as to what spe-
cific things are represented by these symbols. Although the selection of concepts is not subject to any a priori restrictions, it is not arbitrary. The utility of a concept is found in the statements that can be made about them. We select concepts when they can be related to other concepts in a meaningful way. A concept that does not appear to be associated with anything else is of no use whatsoever.

Concepts are justified by their inclusion in laws. Laws, in turn, are statements relating concepts. More precisely, laws describe invariate and observable associations between two or more concepts. If the association is regular but not perfect, we call it "law-like." "If A, then B" is the logical form of a law, but, although this implies necessity, it does not justify any given causal inference. Does A cause B, B cause A, or are both the result of some unidentified C? Perhaps A does cause B, but only when certain other conditions are fulfilled. The problem is that the same observations can support any of the above inferences about the actual causal relationship.

This does not mean that laws are not useful. Oran Young points out that U.S. policymakers would find it very valu-
able to know that Soviet leaders always adhered to their treaty obligations, even if we did not know the causes of this consistent behavior. Knowledge of recurring associations can be quite useful, but such knowledge does not go far enough. Unless we know what forces are actually causal, we cannot be sure that an observed law will continue to obtain in the future. For example, the law that "if the Soviets sign a treaty, then they will abide by its provisions completely" might prove grossly misleading if the conditions that produced this behavior were to change (e.g., a major change in the composition of the Soviet ruling elite). Laws are useful, but unless we have an explanation of the causes that produce them, we cannot be sure of our predictions, and we cannot reliably use our knowledge to control events.

These problems of causal interpretation are particularly acute in the analysis of political events in general and the use of force in particular. First, the method of controlled experimentation is prohibited by the social costs involved. Second, it is often impossible to obtain direct evidence of causal impact, particularly when perceptions and intentions are involved. Firsthand reports by relevant decisionmakers are often unavailable, and the analyst is forced to infer causality from associated changes in
different variables. Third, the use of the armed forces for political purposes rarely, if ever, occurs in isolation. As a result, separating the effects of military forces from the effects produced by the use of other political instruments may be quite difficult. Finally, it is more than likely that the armed forces frequently operate in a symbolic fashion, and their impact is therefore affected by the overall capabilities of the state employing them. Measuring the causal impact of the use of force must account for all these difficulties.

These problems help point the way to a solution. The answer is not to discover more and more associations, or simply corroborate previous findings with additional cases. On the contrary, because the causal interpretation of observed laws is problematic, the appropriate step is to explain the findings, to attempt to show why the observed associations occur.

The explanation of laws is the task of theory. Theories perform this function by showing how the observed associations (laws) are deducible from more general and necessary relations. As Abraham Kaplan has described:
A theory explains because it is more than a mere aggregation of laws; it is a generalization of higher level, the law itself is shown to be a special case of a more general phenomenon. This presents a problem. Although laws are explained and causal relations identified by showing how they can be deduced from a more general set of propositions, we then face the question of why the propositions in the theory can be considered valid. Successful scientific theories resolve this dilemma by making theoretical assumptions. These assumptions are simply stated as true, in order to provide a general picture of the causal relations that govern a given set of observed phenomena. Newton assumed that mass concentrated at a point, among other things; economists assume that man is a rational maximizer, and in international politics, we may assume that states are unitary actors who seek survival. The assumptions are made in order to permit the creation of a theory. If the theory explains observed laws, the merit of the assumptions is revealed.* In addition, theories are judged according to

*It is important to remember that theories are ultimately a product of human creativity. The collection of laws is simply a process of observation; while explaining them involves an indescribable act of insight. The creation of explanatory theories requires the recognition of order amid chaos, and the ability to state what the fundamental relations at work really are.
their explanatory power. We prefer a theory that explains a wide range of phenomena, and shows that events that were previously seen as unrelated are actually "parts of a coherent whole."6

Testing Theories
Because theories are statements of necessary causal relationships that are more general than the laws they explain, they can be tested in a variety of ways. We test theories by inferring hypotheses from them (i.e., by asking what observable consequences should occur if the causal propositions in the theory are true) and by observing if our predictions are confirmed. The greater the variety of tests passed, the more confidence we have in the validity of the theory. Our confidence in the theory is further enhanced when it passes difficult tests. Do the theoretical predictions hold when other factors are believed to work against the causal forces described by the theory? If so, our confidence is greatly increased. We should also seek evidence that helps us choose between competing explanations of a set of observations, by looking "for those consequences of (one) theory whose negation is implied by the alternatives."7 Finally, a failure to pass a test does not lead us to reject the theory. The sensible response is to ask why the theory failed, and if
possible, revise the theory to account for the anomaly. Theories are rejected only when a better alternative is available. 8

The above points are largely common-sensical, but their importance does not seem to be appreciated by most political scientists. Despite the necessity of providing laws with explanatory (that is to say, theoretical) support, many political scientists continue to accumulate laws without asking why they occur, or fail to test the explanations that are proposed in a convincing way. As a result, although suggestive findings may be discovered, their validity is unclear and their value uncertain.

Requirements for Sound Theoretical Work
What is the solution to the difficulties identified above? Good analytical work requires discipline and creativity, and cannot be guaranteed by adherence to a particular method. However, one can at least make such work possible by observing some basic principles. What are they?

First, analysis should proceed with a precisely defined objective in mind. What is puzzling us? What events have attracted our attention? By identifying the problem carefully, we immediately gain a simplification that suggests
what factors are potentially important and what factors are trivial. Second, we collect information according to an at least intuitive sense of what factors may be involved in causing the phenomena under study. Third, as a pattern emerges, we try to explain it by proposing more general hypotheses. This is a creative step, but a necessary one in the development of theories that justify causal inferences. Fourth, the proposed theory must be tested in the manner described above. If a theory correctly identifies the causal relations that produce a set of observations, then these same causal relations should produce similar effects in different circumstances. Testing the theory involves inferring a variety of hypotheses, and seeking evidence to confirm or disconfirm the predictions. Finally, unless the first proposal is completely successful, the research process repeats itself, as information that can resolve anomalies and lead to revised theories is sought and pondered.

Implications
I mentioned a number of distinct difficulties that hinder the analysis of military force as a political instrument above. The conception of law and theory developed here, and the outline of the requirements for sound theoretical
work presented above suggest promising strategies for dealing with these difficulties.

Because controlled experimentation is impossible, we are limited to the historical record. This limitation is serious enough; we should therefore be careful to formulate and test theories through the use of a wide range of historical evidence. Second, although direct evidence is often missing, convincing firsthand accounts are occasionally available, and should be used where possible. Third, and more important, because the use of the armed forces is accompanied by the use of other policy instruments, we should compare cases where force is used with similar events where the armed forces were not employed. A marked difference in the results obtained would suggest the unique causal impact produced by the military option. Finally, although possible excluded variables cannot be controlled by experimentation, the perturbing effects can be minimized by examining cases where force was used in widely varying circumstances. This further suggests the importance of examining cases from different historical periods, and involving different states. If the effects of the political use of military force clearly differ when the cases vary in this manner, it suggests that important variables have been excluded. Most important, the find-
ings must be explained, and the explanations tested by the careful consideration of additional evidence. Theories perform the explanatory task and tell us what the causal relations involved really are. If an explanation is not provided, then factors that are "associated" with a given outcome cannot command an uncritical acceptance. Failure to identify causality is inadequate, because we do not know how to produce the outcomes we want, nor can we estimate all the effects of the instruments we choose. Studies that do not propose and test causal relationships can provide only correlations of uncertain meaning and questionable import.

SUMMARY
The central theme of this section has been that causal inferences are justified when the reasons for an observed association are explained by a more general theory. As a result, analyzing the use of the armed forces as a political instrument requires more than the description of events and accumulation of laws, but must include an attempt to explain the empirical results in a rigorous and consistent way. Unless the explanatory task is accomplished, the findings can only be suggestive, and the precise nature of the causal relations involved remains unknown. We cannot be sure the observed patterns will con-
continue, and we cannot hope to control outcomes reliably through the manipulation of available policy instruments. With these precepts in mind, we can now examine the methods and findings presented in Force Without War.

FORCE WITHOUT WAR: FINDINGS WITHOUT FORCE

In this section, I shall try to show how insufficient attention to the principles of causal inference and theory construction developed above can have serious consequences for the quality of social science work in general and international relations theory in particular. To accomplish this task, I shall examine the methods used in a recent study, Force Without War. I believe that this study contains serious methodological flaws, and that these problems seriously compromise the validity of the findings. A detailed critique of this book will serve to emphasize the importance of understanding how one's substantive interest imposes requirements in theory and for the selection of appropriate methods to satisfy these requirements.

Objectives and Methods of Force Without War

Force Without War is intended as a comprehensive analysis of the role military force has played in recent U.S. for-
eign policy. The principal authors of the study, Barry M. Blechman and Stephen S. Kaplan, have two quite ambitious objectives:

First, to describe the historical record; that is, to identify the incidents in which the United States has used its armed forces for political objectives since the Second World War, and to determine the broad trends in such uses of the armed forces in terms of the context in which the military units were employed and variations in the size, type, and activities of the military units themselves.

Second, to evaluate the effectiveness of the armed forces as a political instrument, in the short term and over a longer period, by analyzing the consequences of such factors as: the size, type, and activity of military units involved in the incidents, the nature of the situation at which they were directed; the character of U.S. objectives; the international and domestic context in which the incident occurred; and the extent and type of diplomatic activity that accompanied the use of the armed forces.

As will be seen later, these are two quite different objectives and the differences should be kept in mind.

In order to "describe the historical record," the study identifies 215 incidents in which U.S. military forces were used for a political objective. The incidents were chosen as those cases that fit the following definition:

A political use of the armed forces occurs when physical actions are taken by one or more components of the uniformed military services as part of a deliberate attempt by the national authorities to influence, or to be prepared to influence, specific behavior of individuals in another nation without engaging in a continuing contest of violence.
The 215 incidents occurring between 1945 and 1975 are then analyzed according to time period, geographic region, political situation (inter- or intra-national conflict), and the participation of other actors. The reader is provided statistical summaries of the types of forces used, the level of force employed, and the various activities performed by U.S. military forces during the 215 incidents. These descriptive statistics reveal a variety of different tendencies. Some of these findings are not surprising (e.g., naval forces are used quite frequently) but some are rather surprising and intriguing (e.g., naval forces performed "traditional" naval missions relatively rarely).

This type of historical summary requires an important caveat. Attempting to predict future behavior from historical trends requires the assumption that what has happened in the past will continue in the future. Is such an assumption warranted? Apparently not completely, at any rate. We are told, for example, that the U.S. is "not likely to engage again easily in Southeast Asia" despite the fact that 28 percent of prior involvements have been in this region, and that the percentage was 36 percent for the period 1966-1975.11 The reasons for this belief are
obvious, but this one example indicates the problems involved in attempting to infer likely behavior from past trends. On the other hand, even if the locus of activity changes, e.g., no more engagements in Southeast Asia, the patterns of activity, instrumentalities, etc., may continue unchanged. As we noted above, however, unless the causes of a set of observations are explained and remain constant through time, we cannot be confident that an observed tendency will continue.

The more important task of this study is the attempt to evaluate the effectiveness of military force as a political instrument. The question of utility is certainly an important one from a policy standpoint, and as such requires careful handling. In order to evaluate utility, a 15 percent sample (33 incidents) is selected from the larger population of 215. The sample is stratified with respect to four variables (level of force, context of U.S. involvement, involvement of other superpowers, and time period) in order to represent the larger population. For each incident, the "operational objectives" desired by U.S. decisionmakers were identified, defined as "the behavior desired of the individual foreign actors in the incident." The decision to focus on operational objectives (as opposed to underlying motivations or overall strategic
goals) was made to "allow the consideration of more objectively determined and empirically verifiable phenomena, without risking the validity of the study's results." Once again, this methodological decision has important consequences, to be addressed below.

The measurement of utility is made by counting the percentage of operational objectives that were achieved in each incident. If at least two-thirds of U.S. objectives were attained, the outcome of the incident was scored as "favorable." This determination was made at two intervals, once six months following the incident and again after a period of three years. In this way, the longevity of a given result can be determined. The number of favorable or unfavorable outcomes can then be compared with changes in various characteristics of the sample incidents.

The study identifies a number of different factors believed to be important in determining the likelihood of favorable outcomes. The conditions examined include:
1. Mode of American action (to assure, deter, compel, or induce)
2. Level of force employed
3. Substantive objectives
4. Style of use of force (direct, indirect, or latent)
5. Type of forces used (naval, ground troops, aircraft, etc.)

In addition, the authors examine the associations between different "situational factors" and the percentage of favorable outcomes. The different situational factors are:

1. Previous uses of U.S. Armed Forces in the region
2. Prior change in the size of U.S. forces deployed in the region
3. Commitments
4. Administration attention and use of coercive rhetoric
5. Personal diplomacy
6. Presidential popularity
7. The role of the Soviet Union
8. Conflict and cooperation between the superpowers
9. The strategic weapons "balance"
10. Structure of the situation (U.S. involvement in inter- or intra-state conflicts)

Changes in these various "internal" and "situational" factors are then compared with the percentage of favorable outcomes. The potential impact of the various factors is assessed by observing whether changes in the various factors were associated with a decrease or increase in the number of favorable outcomes achieved.

The numerous findings that emerge from this extensive aggregate analysis cannot be quickly summarized. A few ex-
amples will suffice to illustrate the character of the results obtained:

1. Favorable outcomes are less frequent when greater levels of force were used, unless nuclear-capable components were included with one or more major conventional force units (i.e., a carrier task force).
2. In the short term, 75 percent of the incidents ended favorably with respect to U.S. operational objectives.
3. In the longer term, the percentage of favorable outcomes declined to less than 50 percent.
4. Outcomes were more frequently positive when the size of U.S. forces in the region had recently changed, either increasing or decreasing.

To their credit, the authors are quite cautious in interpreting their results. As they explicitly point out:

Establishing causality is too heavy a burden to place on our data, and we therefore leave that task to the case studies. Rather, in these chapters we measure a less complicated phenomenon, the extent to which the use of force was associated, in aggregate, with positive or negative outcomes related to the behavior desired of other actors in the sample incidents.\textsuperscript{13}

Recognizing that rough correlations do not provide sufficient grounds for causal inferences, the authors commissioned five regional experts to perform a set of 10 case studies. The case studies examine the following incidents:
1. The Laotian Crisis of 1961
2. Indo-Pakistani War of 1971
3. The Lebanon Intervention of 1958
4. The Jordan Crisis, 1970
5. The Berlin Crisis, 1958
6. The Berlin Crisis, 1961
7. Yugoslavia, 1951
8. The Invasion of Czechoslovakia, 1968
9. The Dominican Intervention, 1961-1966 (viewed as two incidents)

Although the specific case interpretations can be debated, the accounts are detailed and suggestive. Unfortunately, the case studies are not used as profitably as might be expected, for reasons I shall discuss shortly.

The main conclusion of *Force Without War* is that "discrete uses of the armed forces are often an effective way of achieving near-term foreign policy objectives." Although the "success rate" erodes over time, they conclude that the use of the armed forces is justified by the way that it can delay developments, and thereby "provide an opportunity for diplomacy." Their analysis also indicates that short-term success is clearly more frequent than long-term, that reinforcing existing behavior is easier than modifying it, and that the involvement of the Soviet Union works to decrease the percentage of favorable outcomes. One policy recommendation is quite explicit: "the armed forces are an important political instrument. This
role should receive close attention in force planning and operational decisions.\textsuperscript{15}

What value should we place on these findings? How reliable are they? Some of the results are certainly consistent with the conclusions of earlier studies (e.g., that deterrence is easier than compellence) while others have not been observed previously. On balance, what does this study tell us about the role of military force in contemporary international affairs? More precisely, how does the analysis summarized above help us understand the causal role military forces play in determining international outcomes in situations short of war? Do the authors identify a set of causal relations that explain the various findings in a coherent way, and that are capable of further testing and refinement? Does this study permit us to accurately assess the utility of military forces when used in a purely political role? Regrettably, the answer to these questions is no. It is now time to consider the reasons for this failure.

Critique
Assessing the merits of any study requires a recognition of its intentions. \textit{Force Without War} pursues two quite
distinct objectives; first, to describe the historical record, and second, to evaluate the effectiveness of military force for political purposes. We know from our earlier discussion that description, no matter how exhaustive, does not provide an explanation of the observed data. Evaluating the effectiveness of military force is not an easy task, and the difficulty is compounded when the questions involved are not clearly understood or stated, and the type of research necessary for an adequate answer is not performed. Evaluating the utility of military forces requires asking a number of precise questions and selecting a methodology that can provide the best answers. The first question, as noted above, is: What is the unique causal impact of the use of the armed forces? Once this is determined, we can evaluate utility by answering additional questions. Can the effects we have identified be produced by more efficient (i.e., less dangerous or costly) means? Are the unique effects of force (those that do not follow from the use of other instruments) on balance, positive or negative?

Discovering the unique causal impact of military forces requires separating out the effects of their use from other related or intervening variables, and any proposed causal relationship must be accompanied by a more general
explanation. If the use of force produces a particular political effect, how does it do so? By identifying the reasons for an observed association between the use of force and particular outcomes (i.e., by providing an observed association with theoretical support), we gain an understanding of the causal principles governing a realm of phenomena.

Force Without War does not devote adequate attention to this requirement. I believe that this failure lies at the heart of the various specific flaws that damage the validity of the findings. Because the authors are trying to perform both a comprehensive historical description and an evaluation of effectiveness, they are unable to devote sufficient time or attention to the more troublesome problems of measurement and theory-construction that are essential to satisfying their second objective.

The authors are certainly aware of many of the difficulties, but no serious attempt is made to correct them. Instead, their solution is to admit that limitations exist, and then proceed along the same defective path. This may serve to protect them from some criticism, but doesn't strengthen the value of the findings at all.
What are the principal weaknesses of the methodology used in *Force Without War*? If we address it exactly in its own terms -- that is, if we generally ignore the limits imposed by the authors' largely bivariate (as opposed to multivariate) approach to analysis -- it is probably easiest to discuss its five major problems one at a time. We should bear in mind that these are for the most part flaws of research design, not interpretation.

The first problem is a failure to distinguish between important and trivial objectives. In order to simplify the evaluation of outcomes, the authors consider the different "operational objectives" in each case to be of equal importance. Obtaining two inconsequential objectives while failing to obtain a third, vitally important one is scored as a favorable outcome, although the validity of such a judgment is clearly questionable. For example, the emplacement of Army units in Vietnam in 1962 apparently helped reduce Viet Cong attacks and increase the South Vietnamese willingness to fight. These two operational objectives were achieved -- in the short term -- but the use of U.S. armed forces did not affect the North Vietnamese support of the insurgents, arguably the vital objec-
tive both then and later. The overall situation in the South continued to decline, yet this incident is scored as "favorable," because two-thirds of the operational objectives were achieved. Moreover, the relative importance of different incidents (cases) is not examined. Each case is assumed to be of equal value, and the percentage of favorable outcomes remains the sole criterion used to measure how "utility" is affected by different conditions or U.S. actions. No reliable evaluation of the "effectiveness of military force" is possible unless some attempt is made to distinguish between a capacity to achieve important objectives as opposed to trivial ones.

The authors acknowledge that different objectives are of varying importance, and that certain incidents may be more crucial than others. Yet while they are aware of the problem, they do not attempt to correct it. Admittedly, developing a reliable weighting scheme would no doubt be very difficult, and may be impossible. While this may explain why it was not attempted, it also suggests that the entire approach used in the aggregate analyses is flawed, perhaps irreparably.

The second problem is one of bias in case selection. It is more subtle than the first, and is again a direct con-
sequence of the methods adopted in the study. Favorable and unfavorable outcomes are determined by examining whether or not the behavior desired of other actors by U.S. national authorities was performed. Yet U.S. operational objectives in a given incident are chosen, in part, according to the perceived prospects for success. Operational objectives, like any other goal, are selected to reflect the policymaker's estimation of what is desirable and feasible. Decisionmakers are unlikely to use the armed forces when such use is obviously inappropriate and likely to fail.* The 215 cases examined in Force Without War therefore reflect not only American interests (roughly measured as operational objectives) but also a basic estimation by U.S. decisionmakers that the use of force promised some meaningful positive gains. It is important to understand the implications of this fact. Force Without War does not tell us how much the use of the armed forces contributes to the accomplishment of American foreign policy goals, but instead provides an indication of how well force accomplishes those objectives for which force is believed to be well suited. The latter concern is much more limited in scope than the former, and re-

*For example, we do not use the armed forces to improve human rights conditions in the Soviet Union, or to facilitate negotiations on trade agreements with OECD states.
flects the fundamentally narrow theoretical scope of the study.

This bias may be even more pernicious when the armed forces are used to enhance the U.S. image or "appearance of power." Decisionmakers may deploy forces to produce an effect they have reason to believe will occur irrespective of any American action. Cases where an operational objective involves "deterrence of Soviet support" or "maintenance of regime authority" are likely candidates for this type of problem, and including them in the estimate of utility probably results in a further bias.16

Once again, the authors were clearly aware of these problems, for they concede that "fundamental objectives enjoy greater durability than operational objectives, and in the longer term are of more significance."17 Furthermore, they acknowledge that "the chosen objective may...reflect only a determination about what is feasible under the circumstances...."18 Unfortunately, the consequences of this realization are not pursued further. Because the ends are chosen by an awareness of available means, selecting cases according to the presence of a particular set of means and then measuring the attainment of the stated ends will
probably overestimate the effectiveness of the instrument whose effects are being studied, in this case, military force.

Recongnizing this bias helps reveal a third and even more serious weakness in research design: the lack of a control group. Although the primary theoretical objective under consideration is the causal impact of the armed forces in a political role, the authors do not examine *any* cases where the use of force was contemplated seriously but ultimately rejected. If the use of the armed forces as a political instrument has unique and important effects that can be produced by no other instruments of policy, these can only be revealed by comparing the cases in the existing study with cases in which force could have been used, but was not.

In a manner that is by now familiar, this limitation is conceded as well:

...the scientific method might have been brought to bear if the sample of incidents selected could have been compared with a group of incidents in which armed forces might have been used as a political instrument, but, for one reason or another, were not. This approach would require that a much larger number of incidents be examined and that the two sets of incidents be comparative -- that is, share a large number of similar characteristics...this approach does not lend itself very well to analyses in which the subjects are individuals and groups in diverse socie-
ties with different values, levels of information, and so forth.\textsuperscript{19}

The reasons given for the failure to employ the scientific method are not convincing. The extra effort required to examine a larger sample was no doubt of some concern, but fewer, more reliable conclusions are surely preferable to a larger number of findings of more questionable validity. Furthermore, the belief that such "controlled comparison" is not easily achieved when the subjects involved vary greatly in values, levels of information, etc., seems dubious. After all, the incidents that are examined vary widely; indeed, the one requirement of a rigorous comparative approach is to measure the effects of the crucial variable (use of armed forces) over a wide range of other conditions. But if the variable whose alleged effects are the major theoretical interest is constantly present, we cannot reliably infer what its unique causal impact might be.

Fourth, because the incidents examined are all examples of the use of armed forces for political purposes by the United States, this study cannot help us distinguish between the causal impact of the armed forces in and of themselves, and their impact as symbols of American power and capabilities in a host of other areas. One cannot
separate the effects of military actions at the scene from
the effects of actions at home, nor can we separate either
from the latent power of the United States that is contin-
ually present. Even when a particular military gesture
seems clearly associated with a sudden change in behavior,
we cannot attribute all of the causal impact to the mili-
tary action, so long as the enormous latent power of the
United States lurks behind the military signals. Because
American political, military, and economic power has been
pre-eminent over the entire historical period covered by
the study, the causal impact of military actions alone
cannot be determined by the methods used here.

Of course, it can be argued that the study's concern was
to identify the utility of American forces in the contem-
porary era, and that this task is adequately accomplished
by their focus on U.S. incidents. However, because the
utility of force is not compared with that of other in-
struments, and because they do not distinguish between the
effects of force qua force and their role as symbols of
overall superpower capabilities, this study cannot effec-
tively recommend when or why force should be used, or what
its effectiveness will be should American preponderence
decline in other areas.50
Finally, even if we accept the aggregate methodology as given, the authors are occasionally careless in interpreting the data. For example, they observe that favorable outcomes were obtained in all nine cases where land-based combat aircraft were employed. Because this "success rate" is greater than that achieved in the cases where naval forces were used, they infer that this evidence "suggests the greater utility of land-based forces as compared with sea-based forces," and they further suggest that "the Air Force might be used more frequently in political-military operations than has been the case in the past."²¹ Looking at the evidence more carefully suggests that the inference is questionable at best. If we compare the cases where the different types of forces were used by themselves (thereby controlling for possible perturbing or intervening effects when land-based aircraft and sea-based forces were used together), the rate of success is identical (100 percent) and we have more evidence regarding the effectiveness of naval forces. As the following table shows:
<table>
<thead>
<tr>
<th>Types of forces</th>
<th>Favorable outcome total no. of times used alone</th>
<th>Objectives achieved/total no. of objectives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Land-based combat aircraft</td>
<td>2/2 (Cuba, 1960; Laos, 1973)</td>
<td>5/5</td>
</tr>
<tr>
<td>Naval forces</td>
<td>4/4 (Italy, 1947; Jordan, 1956; Cyprus, 1967; Persian Gulf, 1974)</td>
<td>13/13</td>
</tr>
</tbody>
</table>

Source: Force Without War, pp. 76-82, 96-97.

Given all the reservations already described above, this should not be taken as an argument for the greater utility of sea-based forces. It does indicate that even the evidence available can be used to produce questionable conclusions. The strong suspicion that emerges from all the above arguments is that some of the findings presented in Force Without War are erroneous or misleading, although others are no doubt correct. The problem is that the reader has no way of determining which are valid and which are not. 22

Case Studies

Many of the weaknesses of the aggregate analyses -- particularly those concerning the relative importance of different objectives or incidents and the presence of spurious or intervening variables -- might sensibly be dealt with in the case studies. That is what the authors clear-
ly intend, but the detailed case studies also fall victim to weaknesses in research design and implementation. In particular, the unsystematic selection of cases and the reluctance to search for rigorous tests of emerging hypotheses diminishes their utility.

First, it is not at all clear what prior theoretical understanding governed the selection of cases for detailed study. No existing theory is discussed, and why the 10 cases selected warrant detailed examination and comparison is not explained. The reader is told that "cases were selected to highlight variations in the particular circumstances in which the United States has used its armed forces since 1945," but there is more to effective controlled comparison than simply maximizing variation. Cases should be selected with an eye towards maximizing similarities and differences. After all, if everything varies, it is impossible to determine what the effects of any single factor are.

A basic requirement for a controlled comparison approach to case study is that the cases selected must be part of a theoretically significant class of phenomena. As Alexander George has pointed out:
The universe from which cases are selected must be well-defined. This can be done only by stimulating clearly the type of phenomena/event/behavior that is to be studied.

Although the 10 cases examined in *Force Without War* are all examples of the "political use of the armed forces," this class of events is so broad as to introduce enormous problems into any attempt at meaningful comparison. For example, the fact that the specific political objectives sought in each case vary enormously suggests that viewing them as members of the same class may mask important differences. While widely varying cases should be used to test a proposed causal relationship, attempting to discover what causes result from a particular variable requires controlling for the effects of other factors, like the effects of different objectives. To mention an obvious example, in what way is the attempt to assure NATO allies of a continued U.S. commitment following the invasion of Czechoslovakia really comparable to the enforced imposition of a new regime on the Dominican Republic, an act that involved the occupation of the capital by 14,000 U.S. troops? Although 10 cases are examined, the degree to which they can be profitably compared is undermined by the failure to select cases that are representatives of a well-defined, theoretically interesting class.
Second, a notable virtue of case studies is that cases can be selected for which adequate information is available to resolve ambiguities. Although the authors maintain that the task of establishing causality will be left to the case studies, this is not accomplished. For example, Philip Windsor's analysis of the U.S. response to the Czech crisis of 1968 is unable to demonstrate that the American actions had any appreciable effect, and he can only speculate that "U.S. actions deterred the Soviets from actions they might otherwise have contemplated." Given the large number of cases that might have been studied, it is surprising that more attention was not devoted to examining those for which adequate direct evidence of intentions and perceptions was available.

Third, given the wealth of findings produced in the aggregate analyses, one would expect the case studies to test the more interesting, surprising, or problematic findings in a more rigorous way. Unfortunately, this type of careful integration of aggregate analysis and intensive case study is not attempted.
Finally, the chief proponents of the comparative case study format as a valuable research methodology have clearly pointed out that it is best employed in an iterative manner. Cases should be selected according to available information and the insights gained from existing theory. Preliminary findings should prompt re-examination. Further attempt should then be made to resolve theoretical problems through the study of other carefully chosen cases. Have potentially crucial causal factors been omitted? Should different cases be examined? These concerns are not adequately considered in *Force Without War*, although the limited nature of the conclusions certainly suggests that further work was called for. Indeed, the amount of attention devoted to synthesizing the findings and drawing conclusions is slight, especially when compared with other works on similar subjects that have employed this approach.

The Need for Explanation

A final weakness should be addressed, although it is not a direct consequence of the methods used in *Force Without War*. The authors do not attempt to explain their findings in a clear and consistent way. For example, one of the more interesting findings is the relatively minor impact
that changes in the strategic balance between the superpowers appear to have had upon outcomes. Why is this so? The authors mention in passing that this finding may be misleading, because the decline in U.S. strategic preponderance may have prompted U.S. policymakers to scale down their objectives, and therefore the same rate of success is observed over time.

These competing hypotheses, i.e., that the strategic balance has no effect versus the conjecture that its effect is so strong as to have changed the types of U.S. objectives sought, might be resolved through consideration of appropriate case studies suitably handled. In fact, a potentially instructive case is available in the study itself: U.S. objectives toward West Berlin did not change between 1958 and 1961, but the perceived strategic balance shifted from fear of an imminent Soviet advantage to complete confidence in U.S. superiority. Despite this development, the Soviets were more successful in 1961, obtaining tacit U.S. acceptance of the Berlin Wall and the resulting cessation of widespread emigration to the West. Surprisingly, this potential test does not emerge in the case study of the Berlin Crisis or in the later conclusions.
In some cases, to be sure, an explanation is offered in a plausible manner. The high percentage of favorable outcomes associated with the use of land-based combat aircraft is explained by suggesting that such forces indicate a higher level of commitment than naval forces maneuvering offshore. We have shown earlier that the percentage of favorable outcomes properly associated with land-based aircraft is actually not higher than that for naval forces. But, in any event, this sort of explanation should prompt further questions of a more general nature. Why is commitment so important? How is it communicated? This type of question immediately moves the analysis away from the purely empirical level and toward a more general conceptualization of the issues involved in interstate influence. Explicitly abstract, political concepts like motivation, will, interest, etc., are notably absent from this study, largely because the authors apparently prefer to devote the bulk of their attention to an empirical description of trends and tendencies. This omission is striking in view of the degree to which factors like asymmetry of motivation, possession of the status quo, etc., have been identified as crucial in a number of other important studies.
The failure to seek convincing general explanations of the wealth of findings produced in the aggregate analyses helps ensure that the study's contributions to our understanding of how forces affect political events will be minimal. To the extent that a more general explanation is attempted, it is remarkably vague and tentative: "the discrete demonstration of U.S. military capability," when inserted into a situation of "uncertainties and distinct psychological unease...can have a stabilizing effect." This is not an explanation that should satisfy us, because we are left wondering about the causal mechanisms at work, and the conditions necessary for the effect to be produced. They maintain that the military threat must be "credible," but the requirements for "credibility" are not defined. Must the military threat be credible, or can it simply signal possible non-military sanctions?

The proffered discussion of the "special role of force" is really not much more than a sketch of some possible (and, we should acknowledge, plausible) hypotheses. For example: "Establishment of a U.S. military presence may furnish an incentive to a foreign leader for considering the wishes of U.S. policymakers." Or, "Military demonstra-
tions also can ease domestic political pressures...from
groups demanding more decisive action." Both of these
statements may be true, but the converse of each is also
quite possible. An American military demonstration may
decrease the ability of a foreign leader to comply, by
raising his fears that he will appear weak and easily ma-
nipulated. Unless the authors can show why their explana-
tions are to be preferred, the theoretical (and, indeed,
the practical) contributions of the work remain minimal.
Even if we ignore the difficulties already discussed, we
are left with the feeling that the conclusions reached in
Force Without War are the type of hypotheses one employs
at the beginning of a research process, not at the end.
Carefully chosen cases could be used to explore the valid-
ity of the various hypotheses, and a creative search for a
general and "policy-relevant" theory that avoids the in-
herent problems identified above could be undertaken.

SUMMARY
Of course, the task of analyzing the effects of the armed
forces as political instruments is not an easy one. Some
uncertainty and explanatory incompleteness is no doubt in-
evitable, but it should be minimized when possible. It is
the scholar's responsibility to confront the problems and
attempt to find solutions. Admitting that problems exist is easy, but solving them is the important task. The authors of *Force Without War* recognize the weaknesses of their approach and the limitations that these weaknesses impose on their conclusions, but they do not attempt to find solutions, nor do they confront the possibility that the problems may be insurmountable. As a result, the policy implications that they draw from the analysis are of minimal value. Although they advocate that the "political role of the armed forces" should be considered when the requirements for U.S. force structure are determined, it is by no means obvious that the armed forces need to be designed for such a mission. If, as I have suggested, armed forces serve a primarily symbolic function, then their design and procurement should reflect military concerns first, and possible political roles to a far lesser extent.

Why does this extensive and empirically detailed study achieve such minimal results? The difficulty of the questions involved is partially responsible, but I believe that the incompatibility of the two central objectives is also to blame. The methods used to describe the historical record are less than optimal for dealing with the theoretical (i.e., causal) issues that must be addressed
in order to adequately evaluate the effectiveness of the armed forces as a political instrument. This is especially true when adherence to empirical description and generation of extensive "findings" subtracts from the time devoted to rigorous analysis and explanation.

*Force Without War* is an ambitious work, and this is part of the problem. Ambitious goals require considerable effort -- both empirical and theoretical -- in order to achieve meaningful results. As cannot be too strongly emphasized, the examination of more and more fundamentally similar incidents is not profitable unless laws can be discovered and then explained with reference to a more general theory or set of theories whose propositions can be and are rigorously tested.

The authors neglect the requirements of theory construction and devote their primary effort to a quantitative summary of the historical record and the generation of a large number of law-like associations. This effort should not be disparaged, but the attention devoted to data collection was apparently achieved at the expense of analysis of the more important questions. The key point is that the approach used to describe the historical record (the aggregate analyses) is inadequate for the task of deter-
mining effectiveness. The case studies are more promising, but their selection is too haphazard to provide valuable theoretical insights. In short, while the methodology of Force Without War may be well suited to a concise summary of the historical record and the discovery of an infinite number of possibly interesting laws, it cannot, by itself, tell us how effective the use of armed forces has been.

A Final Word

The book is limited in a final sense. In order to maintain "objectivity" and permit the statistical analyses to proceed in a feasible way, no attempt was made to examine the propriety of the goals selected by U.S. policymakers. Admittedly, this is a separate issue, and the authors carefully note that this was not their concern. However, one cannot really evaluate the effectiveness of the U.S. armed forces without eventually asking whether or not the effects of their use have been beneficial or detrimental to American interests. In particular, Force Without War does not devote any serious attention to the potential costs of using force, even apart from the economic burdens. The use of force for limited objectives can create adversaries where none existed before, extend commitments beyond the original intentions of the policymakers initia-
ting its use, and provoke increases in the military capabilities of potential foes. While measuring the costs of the political use of the armed forces was not an explicit goal of the authors of Force Without War, omitting any consideration of this factor further weakens the conclusions that can be drawn from the study.

Much of this review has been quite critical, and a number of virtues of Force Without War have not been singled out. Attempting to provide both a complete record of U.S. political-military responses and a creative explanatory theory that can help us understand the utility of military force is a quite ambitious task. We should perhaps be lenient, recognizing that the problems involved are scarcely trivial. Yet good theoretical work is so rare in our discipline, and "suggestive, but inconclusive" analyses so common, that persistent methodological errors should not be ignored, particularly when the subject matter is of crucial interest to national policymakers. Evaluating the utility of force (i.e., identifying the unique causal impact of particular military actions and developing measures of the positive and negative aspects of the resulting effects) cannot be accomplished by more and more data-gathering, unless the proper information is obtained and is analyzed in a manner that produces clear and general
explanations. The merit of good theory lies in the way it provides more reliable and complete understanding of the causal relations that produce outcomes. This should be our goal, if we are to proceed beyond description.

The authors perform the descriptive task reasonably well, but fail to satisfy the requirements of their second objective. As a result, *Force Without War* consists largely of findings without force. It is a comprehensive catalogue, but provides little theoretical understanding or practical guidance.
FOOTNOTES


2. See Appendix II, available upon request, for an explanation of the symbolic role of military force in the relatively stable bipolar world we have enjoyed since 1945.

3. Karl Popper provides a telling story:

   If I am ordered: "Record what you are now experiencing" I shall scarcely know how to obey this ambiguous order. Am I to report that I am writing; that I hear a bell ringing, a newsboy shouting, a loudspeaker droning, or am I to report, perhaps, that these noises irritate me?


10. Ibid., p. 12.

11. Ibid., p. 32.

12. Ibid., p. 65.

13. Ibid., p. 58.


15. Ibid., p. 534.

16. I am indebted to Kenneth Oye for discussion on this point.

17. Blechman and Kaplan, op.cit., p. 66.

18. Ibid.

19. Ibid., p. 67.

20. One may, of course, argue further that U.S. policymakers are less concerned with why a given policy instrument works than with the fact that it seems to have been "associated" with favorable outcomes in the past. Force Without War might then be seen as satisfying a desire for this kind of "seat-of-the-pants" summary of the results of previous political-military operations. Three problems remain, however. First, as noted above, if the military instrument is not compared with other policy options, policymakers cannot use Force Without War to identify situations where force is unnecessary or inappropriate, or those where the use of force is the best alternative. Uncritical reliance upon past behavior and historical analogies is a well-known danger in policymaking. Because Force Without War does not explain why force worked in some cases but not in others, a policymaker cannot use the study to distinguish appropriate uses of force from foolhardy ones. Second, because the symbolic use of force is closely linked to the overall capabilities of the states involved (see Appendix II), a decline in
American preponderance in other areas will probably decrease the efficacy of our military forces in political roles. Somewhat paradoxically, this may call for more an economic and political retrenchment than for an increase in military capability. Third, because no attempt is made to evaluate the importance of U.S. objectives or accomplishments, we cannot know if the use of force has, over the long run, been beneficial to American interests. We know that the political use of the armed forces has been rather frequent, but we don't know if we are better off as a result. Even the "seat-of-the-pants" policymaker is forced to provide his own evaluation of the utility of American political-military actions since 1945.


22. I am indebted to Jim Thomason for discussion on this point.


26. See below, p. 17.

27. Alexander George's work on coercive diplomacy found that asymmetry of motivation was a particularly crucial variable; and one of the cases examined in his work -- Vietnam -- clearly illustrates how a failure to consider relative motivations can lead to erroneous predictions regarding the efficacy of military force for political objectives. See George, Hall, and Simmons, op.cit., esp. Chapter V. George and Smoke's later work on deterrence explicitly stated that "effective deterrence and signaling require in the first instance that the interests of the U.S. be sufficiently engaged by what is at stake in the area or country in question." George and Smoke, op.cit., p. 560.

Finally, in a remarkable synthesis of case studies and diverse theoretical approaches, Glenn Snyder and Paul Diesing employed formal models of bargaining with theories of decisionmaking and international system structure to investigate how political-military actions can alter the payoff structures held by differ-
ent actors, and thereby affect typical crisis outcomes. Among their more striking findings is that successful resolution of international crises is encouraged when both parties stand firm long enough to permit their relative motivations to be clearly understood by each other. We should note further that this somewhat surprising finding is explained. As the authors state: premature conciliatory strategies actually delay "resolution...because the recipient of the offer is not yet clear about the power relations, and thinks he can get more." A variety of historical evidence is used to support this point. Snyder and Diesing also identify how changes in governmental organization and international system structures affect such propensities as well, which enables their emerging theory to provide explanations over a wide range of historical periods in a manner that is applicable to states other than those examined in the study. Snyder and Diesing, op. cit., p. 489.


29. Ibid., pp. 520-521.
CNA Professional Papers – 1976 to Present*

*Research supported by the National Science Foundation


*Research supported by the National Science Foundation

*Research supported by the National Science Foundation


PP 160: Wilson, Donald P., Jr. The U.S. Sixth Fleet and the Conventional Defense of Europe. 50 pp. Sep 1976. AD A030 417


* CNA Professional Papers with an AD number may be obtained from the National Technical Information Service, U.S. Department of Commerce, Springfield, Virginia 22151. Other papers are available from the Management Information Office, Center for Naval Analyses, 2000 North Beauregard Street, Alexandria, Virginia 22311. An Index of Selected Publications is also available on request. The index includes a Listing of Professional Papers, with abstracts, issued from 1969 to February 1976.