

SUCCESS AND FAILURE IN FOREIGN POLICY

David A. Baldwin

Institute of War & Peace Studies, Columbia University, New York, New York 10027;
email: dab12@columbia.edu

Key Words evaluation, sanctions, force, effectiveness, utility, costs, statecraft, victory

■ **Abstract** The field of foreign policy analysis needs a common set of concepts and analytical frameworks to facilitate comparison of alternative policy options. Not only is general agreement lacking, there is not even a common understanding of what is meant by success. In order to build policy-relevant knowledge concerning success and failure in foreign policy, the following questions must be addressed: How effective is a policy instrument likely to be, with respect to which goals and targets, at what cost, and in comparison with what other policy instruments? Failure to address each question may lead to serious policy mistakes.

INTRODUCTION

Foreign policy is usually viewed as purposive behavior. Specifying the conditions for success or failure of such behavior is arguably one of the most, if not the most, important topic to be studied. Scholarly attention to this topic, however, is not commensurate with its importance. The field of foreign policy studies is preoccupied with the processes of foreign policy making and has tended to neglect the outputs of such processes. In 1975, the *Handbook of Political Science* identified “the concentration on policy process and the neglect of policy output” as “one of the major deficiencies in the study of foreign policy” (Cohen & Scott 1975: 382–83). The situation is not much different today. Most discussions of foreign policy success and failure are left to journalistic pundits or to scholars writing for such journals as *National Interest*, *Foreign Affairs*, *Foreign Policy*, or *The New York Review of Books*.

Although such journals make an important contribution, there is also a need for the kind of rigorous, systematic analysis usually found in more academic journals, such as *World Politics*, *International Organization*, or *American Political Science Review*. Although the research published in such journals is high quality, it is often lacking in policy relevance (Lepgold 1998). The purpose of this article is to formulate an analytical framework for rigorous, systematic, policy-relevant foreign policy evaluation. (For a similar attempt at setting forth a framework for

systematic analysis of foreign policy outputs, see George & Simons 1994 and Baldwin 1985.) Although examples are drawn mostly from the literature on economic sanctions¹ and military force, the approach is intended to be applicable not only to techniques of statecraft, but also to influence attempts of any kind by any actor.

After a brief overview of the state of the field, the article addresses the nature of policy-relevant knowledge and outlines an approach to evaluating foreign policy success. The article also examines the policy context of success and the case of military force. The central purpose is to facilitate scholarly discussion by focusing attention on important concepts and criteria relevant to assessing the success of foreign policy instruments.

FOREIGN POLICY EVALUATION: The State of the Field

The Persian Gulf War in 1991 was preceded by a spirited discussion of the techniques of statecraft most appropriate for dealing with Iraq's invasion of Kuwait (Freedman & Karsh 1993). Some analysts favored exploring every last diplomatic possibility, some advocated reliance on economic sanctions, and others argued that only military force would do the job. Clearly, what was needed was an analytical framework for comparing the utility of various types of statecraft. However, to the extent that scholarly literature on the utility of techniques of statecraft existed, it was compartmentalized into various "islands" of literature. There was a literature on economic sanctions that (endlessly) debated the question of whether such measures work (for references, see Baldwin 1985 and Hufbauer et al 1990) but had little to say about the utility of sanctions in comparison with other techniques of statecraft. There was a literature on military force that discussed its pros and cons but rarely discussed nonmilitary alternatives to force (for references, see Art & Waltz 1999, Shultz et al 1993, Baldwin 1995). The literature on diplomacy was diffuse in focus and rarely attempted to evaluate its utility relative to other policy instruments (e.g. Watson 1983). And the literature on propaganda barely existed in 1990 (for references, see Jowett & O'Donnell 1992 and Lasswell et al 1979/1980).

¹For reasons explained elsewhere (Baldwin 1985), I prefer "economic statecraft" (both the term and the concept) to "economic sanctions." The term economic sanctions is used here for two reasons: (a) the literature on economic sanctions is more easily identifiable than that on economic statecraft, which is a broader concept; and (b) the term economic sanctions is often used loosely to refer to a large part of economic statecraft. By using the term economic sanctions, I reluctantly acquiesce to (sloppy) common usage. Since this grudging acquiescence has recently been confused with advocacy (e.g. Pape 1997, 1998), I wish to make it clear that I neither approve of nor advocate use of the term economic sanctions. The fact that one sins does not mean that one approves of sin, much less that one advocates it.

The conventional wisdom is that economic sanctions do not work, i.e. they have a low rate of success (e.g. Tsebelis 1990, Kunz 1997, Morgan & Schwebach 1997, Pape 1997, Preeg 1999). One quantitative study estimates that they have a success rate of 35% (Hufbauer et al 1990). To describe this success rate as low, however, implies some criterion of judgment, a criterion that is rarely specified. Is 0.350 a low batting average for a baseball player? Only those who know nothing about the game of baseball would agree. The problem is that we know more about the game of baseball than we do about the "game" of foreign policy. Is 0.350 a low "batting average" for a foreign policy instrument? Such knowledge does not exist. It is often implied, however, that competent and knowledgeable policy makers should achieve a much higher success rate when using economic sanctions (e.g. Tsebelis 1990).

Although many studies address the question of whether economic sanctions work, very few address such questions as the following: Does military force work? Does diplomacy work? Does propaganda work? Despite the paucity of such studies, conventional wisdom holds that military force usually works. Pape (1997:90) asserts that "military instruments are often thought to be the only effective means for achieving ambitious foreign policy goals like taking or defending territory."² Art (1996:10) contends that force is "central to statecraft." Even works that downplay the importance of force in international affairs view it as more effective than other means if costs are ignored (Keohane & Nye 1989:16–17; see also Baldwin 1989:151–55).

The literature on foreign policy evaluation is also characterized by analytical and conceptual anarchy. Analytical approaches and conceptual definitions abound. Authors not only disagree as to whether various techniques of statecraft work; they disagree on the very definition of "work."

The chaotic state of the field is illustrated by a recent debate about the success of international peacekeeping operations (Druckman & Stern 1997). Five scholars disagreed with respect to numerous issues, including (a) the meaning of the terms success and failure; (b) the relevance of counterfactuals;³ (c) the relevance of costs; (d) actor designation (e.g. national government, international organization, or humankind); and (e) whether to judge success in terms of actor goals or in terms of higher values, such as global peace and justice. Noting that "the difficulties in evaluating peacekeeping missions are both conceptual and methodological," the monitors of the debate are pessimistic about future progress. "Even though further progress depends on developing a broad conceptual framework that can guide evaluation, given the differences in perspectives evident from the remarks of the experts consulted here, we are unlikely to see research in the near future guided by a single analytical framework" (Druckman & Stern 1997:163–64).

²The Louisiana Purchase (1804) apparently does not qualify as "taking territory."

³The "positivist" scholars expressed the most reservations about counterfactual analysis. Given the importance of causal analysis to most positivists, this is, to say the least, surprising.

Despite such pessimism, the purpose of this article is to propose an analytical framework applicable to judging the success of peacekeeping operations, economic sanctions, military undertakings, and other types of influence attempts.

WHAT IS POLICY-RELEVANT KNOWLEDGE?

Policy-relevant knowledge is what policy makers need to know in order to choose among alternative courses of action. It refers to the rational adaptation of means to ends. Ultimately, policy-relevant knowledge is not a substitute for wisdom, since it is concerned only with intermediate goals, i.e. those that are means to higher ends. Policy analysis, therefore, focuses on the efficient pursuit of given ends.

Wildavsky (1979) has suggested a seemingly contrary view that “creativity” in policy analysis consists of formulating problems that have solutions. He illustrates the point with the following story:

Mike Teitz tells about a soldier in New Zealand who was ordered to build a bridge across a river without enough men or material. He stared along the bank looking glum when a Maori woman came along asking, “Why so sad, soldier?” He explained that he had been given a problem for which there was no solution. Immediately she brightened, saying, “Cheer up! No solution, no problem.” (Wildavsky 1979:3)

Creativity, however, is not necessarily the same as rational problem solving. It may be true that such an approach stimulates creative thinking, but it is not a substitute for rational adaptation of means to ends. A country with an unexpected budget surplus of \$10 billion will have no trouble finding problems to solve. The difficulty lies in choosing among them. Which problems should be solved and which should be addressed at a later time? Of the many alternative ways to spend the money, which will maximize the utility (i.e. welfare) of the country? Problems do not disappear just because solutions are not available or cannot be readily identified. The soldier in the story may have been cheered by the woman’s admonition, but it did not make his problem go away.

A better way to characterize policy analysis is, “No scarcity, no problem!” Policy choice involves choosing among alternative courses of action under conditions of resource scarcity. If everyone can have his cake and eat it too, there are no choices to be made and no need for policy on this issue. Policy problems arise because there are not enough resources available to cope with all of the problems facing policy makers. Under such conditions, policy makers need ways to compare alternative courses of action, i.e. alternative ways of using scarce resources.

Knowledge about the likely utility of a technique of statecraft, therefore, is policy-relevant only insofar as it is helpful in deciding whether or not to use one technique rather than another. Is it helpful to know that economic sanctions only work 35% of the time or that they have a 0.35 probability of working in a particular

case? Without comparable knowledge about the likely utility of alternative policy instruments, such knowledge has no policy relevance whatsoever.

ESTIMATING POLICY SUCCESS

Estimating the success or failure of policy instruments is difficult because the concept of success is slippery, recipes for success can be misleading, the dimensions of success are multiple, and clear-cut victories or defeats are few.

The Concept of Success

Success is a slippery concept. Unlike power or wealth, success is not just one of many goals that people may choose to pursue. To the extent that human behavior is purposeful, everyone may be said to pursue success. This is because success is defined in terms of favorable or desired outcomes. Both the definition of success and the implicit rules used in applying the term suggest that costs are an important part of the concept. Successful undertakings are those without excessive costs. Winning a nuclear war by destroying life as we know it or imposing economic sanctions that secure compliance of the target state only by bankrupting the country that imposed them are unlikely to be described as instances of success. "The operation was a success, but the patient died" does not mean what it seems to say. It is a sardonic expression implying that the term success is being misused. If success is defined in terms of favorable policy outcomes, it is necessary to consider both costs and benefits in assessing the success of an undertaking. The concept of a Pyrrhic victory implies a difference between real and apparent victory, and it is as relevant to nonmilitary as to military statecraft.

Recipes for Success

Some recipes for success in foreign policy making are difficult to reconcile with the above concept of success. One such recipe is the advice to pursue modest goals (e.g. Leyton-Brown 1987:309, Druckman & Stern 1997:157, 163). The poet's admonition that "a man's reach should exceed his grasp" is described by one author as "a fine philosophy of life, but a poor prescription for economic sanctions" (Leyton-Brown 1987:309). Such recipes have more to do with creating the illusion of success than with achieving it.

Logically, the specification of a standard of achievement precedes specification of the determinants of, or conditions for, the success of an undertaking. Any statement that purports to specify the determinants or conditions of success presupposes a concept of success. Take, for example, the following recipe for success: "If you want to get to the other side of this river, you will have to swim or find a bridge." In this example, getting to the other side of the river is the operational definition of success, and swimming or finding a bridge are alternative means for accomplishing this goal. The advice to change your goal from crossing the river to

the more modest one of staying on this side of the river is not a recipe for success; it is a redefinition of success. This bit of semantic sleight-of-hand should not be misconstrued as advice on how to succeed. It makes little sense to describe someone as pursuing success without specifying success in doing what. If a prospective student asks for advice on how to succeed in college, one must first ascertain whether the goal is to get good grades or to get a sound education. "Take easy courses" may be good advice with respect to the first goal, but "take challenging courses" may be more appropriate advice with respect to the second. Likewise, when advising foreign policy makers on how to make economic sanctions successful, one must first ascertain the goals in terms of which success is to be defined. Advising them to pursue modest goals or not to "bite off more than they can chew" violates this basic precept of prior goal specification. Those who never run a race will never lose a race, but they will never win one either. Economic sanctions that are never used will never fail, but they will never succeed either. "Stick to easy things" may be a recipe for avoiding failure, but it is hardly a recipe for success.

Another so-called recipe for success is the "more is better" approach that characterized military planning for the Persian Gulf War. The idea that massive and overwhelming military force is always preferable to graduated escalation was one of the lessons that some military thinkers derived from the Vietnam War (Gacek 1994, Summers 1981). The problem with this recipe is that it fails to provide guidelines for determining how much is enough. More always seems to be better. This perspective is likely to appeal to those who conceive of national security in terms of "interests that are pursued notwithstanding the costs incurred" (Leffler 1990:145). From the standpoint of a rational policy maker, however, there are no such interests.

A more defensible, but less glamorous, recipe is the marginal utility approach, which prescribes using a technique of statecraft until the marginal benefits of doing so equal the marginal costs of doing so. Instead of "never bite off more than you can chew," the marginal utility approach would counsel foreign policy makers to take successively larger bites until the bites became unchewable, then back off a little. Those who never bite off more than they can chew may seldom choke, but they are unlikely to make maximum use of their chewing ability either. "Take small bites" is hardly a recipe for great accomplishments. It also begs the question of how much smaller than the chewable maximum the bite should be.

Dimensions of Success

In business, as in foreign policy making, success is multidimensional. In business, the costs of advertising, marketing, and production must be considered along with the revenues produced by sales in assessing the overall success of a firm. In foreign policy, the effects on allies, the trade-offs among national interests, and the effects on adversaries must all be considered in assessing the overall success of a technique of statecraft. The difference is that business firms have the common denominator of money; they can add up the consequences of their various activities in monetary

terms in order to calculate a “bottom line,” i.e. profits. Foreign policy makers have no such standard of value and must confine themselves to rough judgments in estimating the overall success of an undertaking. It is nevertheless helpful to identify some dimensions of success with respect to which such judgments may be made.⁴

Effectiveness

Since (most) foreign policy is goal-oriented,⁵ evaluating effectiveness in accomplishing goals is a necessary but not sufficient ingredient in estimating success. If there are concepts of success that do not include goal attainment, they have not yet come to the attention of this writer. This is not to say that estimating effectiveness is simple or easy. Consider the following propositions:

1. Foreign policy makers usually pursue multiple goals with respect to multiple targets. During the Persian Gulf crisis, the goal of US foreign policy makers was not solely to force Saddam Hussein to withdraw from Kuwait. Additional goals included restoring the government of Kuwait, minimizing damage to Kuwait, discouraging Israeli intervention, encouraging United Nations support, reassuring potential allies that the United States was determined but not trigger-happy, discouraging other potential aggressors from trying to emulate Iraq's behavior, and so on. These various goals and targets were not equally important, but neither were they trivial enough to justify ignoring them. “Winning the war” is an oversimplification of the goals of any war, and “achieving the primary goal” is an equally misleading way to define the success of economic sanctions. Despite the widespread acknowledgment that the goals and targets of foreign policy tend to be multiple, success is often measured solely in terms of primary goals and targets (see Baldwin 1985).
2. Policy change and behavior are not the same thing. When President Reagan authorized the exchange of weapons for hostages, it was the policy of the United States not to negotiate with terrorists on such matters. Even “behavior” is more complex than a narrow behavioristic focus on overt actions would suggest (on narrow behaviorism, see Oppenheim 1981:191). Dimensions of behavior that are of potential interest to foreign policy makers include not only easily observable policy changes, but also changes in beliefs, attitudes, opinions, expectations, emotions, and/or predispositions to act.
3. Goal attainment is a matter of degree (Simon 1976:177). Art (1996:24) has applied this insight to foreign policy by observing that “a given instrument

⁴Although the dimensions of success discussed here are not arbitrary, neither are they definitive. There are many ways to divide success.

⁵Even those who view foreign policy as expressive rather than instrumental do not dispute that it is usually instrumental.

can carry a state part of the way to a given goal, even though it cannot carry the state all the way there. At one and the same time, an instrument of statecraft can usefully contribute to attaining many goals and yet by itself be insufficient to attain any one of them.” Some writers, however, ignore gradations in effectiveness and classify policy outcomes in terms of either success or failure (e.g. Pape 1997).⁶ Levy (1969:95) labels this classification the fallacy of misplaced dichotomies. “To set up a distinction in binary form when the things referred to vary by degree or in some other fashion is not only the classic misuse of the law of the excluded middle, it also guarantees the begging of important questions” (Levy 1969:95). Dahl (1976:26) labels the same phenomenon “the lump-of-power fallacy” (see also Baldwin 1985:130–31).

The spectrum of degrees of goal attainment should not be confined to positive numbers. The possibility of negative goal attainment should be considered. For example, an attempt to undermine the stability of a regime (e.g. Castro or Saddam) might trigger a rally-round-the-flag effect that strengthens the regime. The possibility of such “negative power” was recognized long ago by Dahl (1957).

Costs to the User

Evaluating the success of a foreign policy instrument solely in terms of effective goal attainment is akin to evaluating the success of a business firm solely on the basis of sales. A firm that sells a million widgets but loses money is less successful than a firm that sells only two widgets but makes money. Dahl & Lindblom (1953:38–39) long ago explained why costs must be considered in estimating success:

An action is rational to the extent that it is “correctly” designed to maximize goal achievement.... Given more than one goal (the usual human situation), an action is rational to the extent that it is correctly designed to maximize net goal achievement.... The more rational action is also the more efficient action. The two terms can be used interchangeably.... An action is “correctly” designed to maximize goal satisfaction to the extent that it is efficient, or in other words to the extent that goal satisfaction exceeds goal cost.⁷

In judging the success of instruments of statecraft, as in judging the success of a business firm, costs should be an important part of the calculation. Unfortunately, the success of economic and military statecraft is often estimated solely on the basis of goal achievement without reference to the costs incurred (e.g. Blechman & Kaplan 1978; Leyton-Brown 1987; Tsebelis 1990; Art 1996; Pape 1996, 1997;

⁶Pape (1996) also ignores degrees of success in assessing the effectiveness of coercion by military means (cf Mueller 1998:204–5).

⁷On success as net value, see also Simon (1976) and Knorr (1966).

Morgan & Schwebach 1997). The following passage provides a typical view of “successful” sanctions:

What does it mean for sanctions to be successful? In general, we are interested in determining the conditions (if any) under which economic sanctions produce an intended (on the part of the sanctioner) change in policy by the target state.... The focus of the debate regarding the effectiveness of sanctions is on whether they can enable the sanctioner to achieve its goals of altering the behavior of the target. (Morgan & Schwebach 1997:29)

Clearly, costs are not part of this conception of success; effectiveness is everything.

Costs to the Target

Inflicting costs for noncompliance on the target of an influence attempt is often used as a measure of success (Dahl 1968, Harsanyi 1962). Other things being equal, the higher the costs for noncompliance that an instrument of statecraft inflicts on the target, the more successful it is. Imposing costs for noncompliance is sometimes confused with failure. Pape (1998:197), for example, objects to treating the costs inflicted for noncompliance as a measure of success: “The fact that a target that refuses to concede may suffer substantial costs does not turn failure into success.” Schelling responds to such reasoning in a brilliant essay entitled “The Strategy of Inflicting Costs.” Is it worthwhile for one’s adversary to spend money on a bullet if one can protect oneself with the purchase of a bulletproof vest? Schelling concludes, “He has wasted his money if the vest is cheap, made a splendid investment if my vest is expensive, and if asked what he accomplished by buying his bullet should have the good sense to say that he imposed a cost on me, not that he hoped to kill me and was frustrated” (1984:274).

Bueno de Mesquita (1981:90) uses logic similar to Schelling’s to explain why even countries that expect to lose a war may find it worthwhile to fight:

What, then, can a victim who expects to lose gain from fighting? Presumably such a victim can hope to impose a sufficient cost on the opponent to reduce the concessions that have to be made at the time of surrender. It may hope, for instance, to impose enough costs to prevent the need for a total, unconditional surrender.

To classify such a situation as total failure would be misleading.

Stakes for the User

Not all foreign policy goals are equally important. Winning World War II and winning the release of a political prisoner in another country are not equivalent accomplishments. To weight such achievements equally in judging the overall success rate of a technique of statecraft would be misleading. Other things being

equal, the bigger the stakes, the more valuable is the degree of achievement and the more successful is the influence attempt.

Stakes for the Target

The more the target has at stake, the more difficult the undertaking is likely to be. [Among the few to recognize this point are Morgan & Schwebach (1995: 259–60) and Kirshner (1997:34); see also Baldwin (1985:133).] In competitive diving, the points awarded for execution are weighted for the difficulty of the dive. A similar scheme is appropriate for judging the success of a foreign policy instrument. Thus, a small degree of goal achievement in a difficult task might constitute a greater success than a higher degree of achievement in an easy task. Deterring the Soviet Union from launching nuclear missiles at the United States was probably a relatively easy task, although the stakes were very high for the United States. By comparison, getting South Africa or Rhodesia to change the way their societies were governed was relatively difficult, even though the stakes were lower for the United States. Other things being equal, the more difficult the undertaking, the more valuable is the achievement.

Other Criteria

Although the criteria specified above seem relevant to any foreign policy undertaking, other criteria may be applicable as well (for discussion, see Dunn 1994:282–89, Dahl & Lindblom 1953). For example, the equity with which the costs and benefits of the undertaking are distributed may be an important consideration in some situations. Economists often use Pareto optimality as a criterion for judging the success of trade negotiations. Some might want to use the criterion of adequacy, the extent to which the influence attempt solves the problem at hand. One could plausibly argue, for example, that the use of military statecraft with respect to Iraq in 1991 produced an adequate solution to the problem even though it did not achieve all of its goals completely.

SUCCESS IN A POLICY CONTEXT

The temptation to infer policy implications from estimates of past or future success of a policy instrument seems irresistible. Consider the following example: “If policy makers are aware that sanctions can rarely have an impact (and they should be) then sanctions should occur only in those instances in which there is a fair chance that they would ‘work’” (Morgan & Schwebach 1997:45–46).

The question of whether a technique of statecraft is likely to work is different from the question of whether it should be used (George & Simons 1994:268–69). Knowledge about the likely success of a foreign policy instrument provides no useful guidance to policy makers as to whether it should be used. Only comparative analysis of the prospective success of alternative instruments provides policy-relevant knowledge. Even if the expected utility of using a technique of statecraft

is low, the expected utility of alternative techniques may be even lower. Before one can agree with Nossal (1994:xv) that “sanctions are a notoriously poor tool of statecraft,” it is necessary to consider alternative policy options. No matter how much one detests swimming, it may seem quite appealing if the only alternative is sinking.

The need for comparative evaluation of both the costs and benefits of policy alternatives is demonstrated by Simon (1976) and Bueno de Mesquita (1981).

An administrative choice is incorrectly posed, then, when it is posed as a choice between possibility A, with low costs and small results, and possibility B, with high costs and large results. For A should be substituted a third possibility C, which would include A plus the alternative activities made possible by the cost difference between A and B. If this is done, the choice resolves itself into a comparison of the results obtainable by the application of fixed resources to the alternative activities B and C. (Simon 1976:179)

Bueno de Mesquita (1981:183) employs similar logic with respect to the alternatives of force and diplomacy:

Leaders expecting a larger net gain through diplomacy than through war ... should rationally elect to pursue their goals through diplomatic bargaining and negotiating. This is true even if the expected gross gain from war is larger than the gross gain from diplomacy, provided that the cost differential is large enough (as it frequently is) to make the net effect of diplomacy preferable to war.

The implication of such reasoning is that one technique of statecraft may be preferable to another even when the former is more likely to achieve a given set of goals, provided the cost differential is big enough.

The purpose of evaluating the success of policy instruments is to judge the wisdom of past or future use of such instruments. In order to make such judgments, one must ask whether some other instrument would have worked (or will work) better. The existence of policy options is sometimes denied. For example, it is often said that in August 1990, at the beginning of the Persian Gulf crisis, the United States had no military options (e.g. Freedman & Karsh 1993:67). Likewise, policy makers sometimes justify past decisions by asserting that they had no alternative. It is the responsibility of policy analysts, however, to point out that policy makers always have options. A policy maker can legitimately claim that he did the best he could in a difficult situation, but he cannot claim there was no other option. “I did the best I could” is a legitimate plea for understanding; “I had no alternative but to do what I did” is an attempt to absolve the policy maker from responsibility for his or her decisions.

Foreign policy instruments are used in situations characterized by strategic interaction. The targets of such influence attempts are not passive or inert objects; they are likely to react. Tsebelis (1990) suggests that both policy analysts and policy makers commit what he calls the Robinson Crusoe fallacy by failing to take account of strategic interaction. Although it is true that the intended effects

of a foreign policy instrument are often offset or nullified by reactions of the target state, the only implication for policy is to counsel foreign policy makers against expecting too much. Because all techniques of statecraft are employed in the context of strategic interaction, one must assume that all are equally affected by such interaction unless there is evidence to the contrary. And if strategic interaction affects all techniques equally, it has no relevance to the policy maker trying to choose among them. Policy makers are only interested in the differences among policy instruments, not in the similarities. Strategic interaction may lower the probability that a given policy instrument will be successful, but as long as some techniques have more probability of success than others, the rational policy maker will continue to use them. Sometimes foreign policy makers are faced with a dismal set of alternatives.

MILITARY FORCE

Military force remains an important instrument of statecraft. Given the potential damage that military statecraft can produce, especially in the nuclear age, it is important to estimate its utility accurately and in ways that facilitate comparison with nonmilitary techniques of statecraft. Unfortunately, neither the costs nor the benefits of military statecraft have received the scholarly attention they deserve.

The outcomes of wars tend to be characterized in dichotomous terms as victory or defeat. Considering the importance of assessing war outcomes, one might expect to find a large literature on the nature of military success. Indeed, one could argue that the central point of the most famous book ever written on war (i.e. Clausewitz 1976) concerned the criteria appropriate for determining success in war. Nevertheless, the literature on military force contains few discussions of the meaning of success (for exceptions, see Jervis 1989:16–19 and Hobbs 1979). Even quantitative studies that are rigorous in many respects rely on “the consensus among acknowledged specialists” (Small & Singer 1982:182; Wang & Ray 1994) in assessing war outcomes. One is left to wonder who these specialists are, how they arrived at their conclusions, and how their consensus was determined.

Although Clausewitz wrote before the invention of game theory, he clearly laid the groundwork for thinking about war in nonzero-sum terms. Despite Clausewitz and despite Schelling’s (1984:269) contention that war itself is “a dramatically nonzero-sum activity,” the tendency to treat war as a zero-sum conflict persists in the literature on military statecraft. War outcomes are usually coded in terms of “win,” “lose,” or “draw,” which are consistent with zero-sum games, rather than in terms of the degree to which each participant was able to achieve its multiple goals (see Art 1996:9, Bueno de Mesquita 1981, Small & Singer 1982, Wang & Ray 1994, Stam 1996). Indeed, in nonzero-sum games, it is conceivable that all participants may be winners or all may be losers. A global nuclear war would make this point painfully obvious.

Military statecraft tends to be the most costly means of pursuing foreign policy goals. Annual defense spending by the United States approaches \$300 billion, while spending on nonmilitary statecraft (State Department, United Nations, etc) is less than \$10 billion. Of course, not all of the costs of statecraft are measurable in terms of government spending. Building nuclear weapons, for example, creates some probability of nuclear war. The costs of such a war, discounted for its likelihood, should be part of the calculus of costs. To argue otherwise would be akin to the contention that playing Russian roulette is not costly if the player wins. If he is using a six-shooter, the correct cost calculation would include the probability of one in six that he will die.

Economic sanctions also have costs that are not measurable in budgetary terms. For example, when trade is suspended, the costs of business foregone should be part of the cost calculus of sanctions, as business groups are continually pointing out. Even so, estimates of the value of business foregone as a result of US economic sanctions do not begin to approach the annual defense budget. Military force is not always more expensive than other policy instruments, but it usually is.

If military force is so expensive, one might expect those who study it to be especially concerned with its costs. This is not necessarily the case. Of course, it is not difficult to identify students of military statecraft who give serious attention to the costs of using it, e.g. Jervis, Stam, Knorr, Schelling, Brodie, George, and Bueno de Mesquita. It is equally easy, however, to find works that purport to say something about the utility of military statecraft but devote little or no attention to the costs of using it (Art & Waltz 1999, Pape 1996, Blechman & Kaplan 1978, Art 1996). And although nuclear weapons have been an important ingredient in military statecraft for the last half century, a recent study by the Brookings Institution (Schwartz 1998) argues that the costs of such weaponry have received woefully inadequate attention.

Cost is an intrinsic part of the concept of utility, and utility calculations are what estimates of success are all about. The concept of a Pyrrhic victory implies that appearances can be deceiving with respect to war outcomes. Costs matter. A Pyrrhic victory is no victory at all. To evaluate the success of a war outcome without reference to the costs incurred by the participants is seriously misleading. Although Small & Singer (1982:182) ignore costs in coding winners and losers of wars, they at least admit that this is a possible weakness in their approach. Other writers (e.g. Wang & Ray 1994) include no such admission.

Clausewitz's (1976:87) description of war as "the continuation of policy by other means" implies that success or failure in war should be assessed in the same way that the success or failure of other policy means are assessed. It implies that military force is simply an alternative instrument available to policy makers for the pursuit of particular ends. If policy makers are to choose between military force and alternative means, they must have an analytical framework that facilitates comparing the likely utility of one instrument with that of others. In assessing war outcomes, as in estimating the success of other instruments, success is best considered a matter of degree; zero-sum assumptions are misleading; and cost considerations are essential.

CONCLUSION

Foreign policy decisions often have momentous consequences. Providing foreign policy makers with the kind of knowledge that would help them to choose more rationally among various instruments of statecraft deserves higher priority among scholars than it has received. The emergence of such knowledge has two requisites.

First, the nature of policy-relevant knowledge must be understood. In order to make rational policy decisions, policy makers need to ask how effective a policy instrument is likely to be, with respect to which goals and targets, at what cost, and in comparison with what other policy instruments. Some have suggested that partial knowledge has policy value. Pape, for example, maintains that “learning that one of the instruments is rarely effective is in itself a finding of ... policy value” (1998:198). But a little knowledge can be a dangerous thing. Failure to ask all of the questions above can lead to serious policy mistakes. This is not to suggest that every scholar must address all of these questions. Any study of the utility of techniques of statecraft that fails to address all of the above questions, however, should carry a disclaimer similar to that on cigarette packages, e.g. “Warning: the knowledge contained in this study omits some of the important questions essential to rational policy making; any attempt to formulate policy based on this study could be hazardous to the health of that policy.”

Second, the development of concepts and analytical criteria that permit the comparison of alternative techniques of statecraft is both possible and desirable. In order to compare military force and economic sanctions, for example, a conception of success common to both and a set of criteria applicable to both are necessary. Any approach that fails to allow for degrees of success (or failure) or fails to account for both the expected costs and benefits of each technique can seriously mislead policy makers.

Before criticizing foreign policy failures, one should ponder the meaning of success and failure in foreign policy. Before reiterating the Vietnam-era slogan of “Why not victory?” one should ask, “What does victory mean, and how would we know it when we see it?”

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

- Art RJ. 1996. American foreign policy and the fungibility of force. *Secur. Stud.* 5:7–42
- Art RJ, Waltz KN, eds. 1999. *The Use of Force: Military Power and International Politics*. New York: Rowman & Littlefield. 5th ed.
- Baldwin DA. 1985. *Economic Statecraft*. Princeton, NJ: Princeton Univ. Press
- Baldwin DA. 1989. *Paradoxes of Power*. New York: Basil Blackwell
- Baldwin DA. 1995. Security studies and the end of the cold war. *World Polit.* 48:117–41
- Blechman BM, Kaplan SS. 1978. *Force Without War: U.S. Armed Forces as a Political Instrument*. Washington, DC: Brookings Inst.

- Bueno de Mesquita B. 1981. *The War Trap*. New Haven, CT: Yale Univ. Press
- Clausewitz C. 1976. *On War*, ed./transl. MH Howard, P Paret. Princeton, NJ: Princeton Univ. Press
- Cohen BC, Scott HA. 1975. Foreign policy. In *Handbook of Political Science*, Vol. 6. *Policies and Policymaking*, ed. FI Greenstein, NW Polsby. Reading, MA: Addison-Wesley
- Dahl RA. 1957. The concept of power. *Behav. Sci.* 2:201–15
- Dahl RA. 1968. Power. *International Encyclopedia of the Social Sciences*, Vol. 12. New York: Free
- Dahl RA. 1976. *Modern Political Analysis*. Englewood Cliffs, NJ: Prentice-Hall. 3rd ed.
- Dahl RA, Lindblom CE. 1953. *Politics, Economics, and Welfare: Planning and Politico-Economic Systems Resolved into Basic Social Processes*. New York: Harper & Row
- Druckman D, Stern PC. 1997. Evaluating peacekeeping missions. *Mershon Int. Stud. Rev.* 41:151–65
- Dunn WN. 1994. *Public Policy Analysis: An Introduction*. Englewood Cliffs, NJ: Prentice-Hall
- Freedman L, Karsh E. 1993. *The Gulf Conflict: 1990–1991*. Princeton, NJ: Princeton Univ. Press
- Gacek CM. 1994. *The Logic of Force: The Dilemma of Limited War in American Foreign Policy*. New York: Columbia Univ. Press
- George AL, Simons WE, eds. 1994. *The Limits of Coercive Diplomacy*. Boulder, CO: Westview. 2nd ed.
- Harsanyi JC. 1962. Measurement of social power, opportunity costs, and the theory of two-person bargaining games. *Behav. Sci.* 7:67–80
- Hobbs R. 1979. *The Myth of Victory: What is Victory in War?* Boulder, CO: Westview
- Hufbauer GC, Schott JJ, Elliott KA. 1990. *Economic Sanctions Reconsidered*. Washington, DC: Inst. Int. Econ. 2nd ed.
- Jervis R. 1989. *On the Meaning of the Nuclear Revolution: Statecraft and the Prospect of Armageddon*. Ithaca, NY: Cornell Univ. Press
- Jowett G, O'Donnell V. 1992. *Propaganda and Persuasion*. London: Sage. 2nd ed.
- Keohane RO, Nye JS. 1989. *Power and Interdependence*. Boston: Scott, Foresman. 2nd ed.
- Kirshner J. 1997. The microfoundations of economic sanctions. *Secur. Stud.* 6:32–64
- Knorr K. 1966. *On the Uses of Military Power in the Nuclear Age*. Princeton, NJ: Princeton Univ. Press
- Kunz DB. 1997. *Butter and Guns: America's Cold War Economic Diplomacy*. New York: Free
- Lasswell HD, Lerner D, Spier H, eds. 1979/1980. *Propaganda and Communication in World History*. 3 vols. Honolulu: Univ. Hawaii Press
- Leffler MP. 1990. National security. *J. Am. Hist.* 77:143–65
- Lepgold J. 1998. Is anyone listening? International relations theory and the problem of policy relevance. *Polit. Sci. Q.* 113:43–62
- Levy MJ. 1969. "Does it matter that he's naked" bawled the child. In *Contending Approaches to International Politics*, ed. K Knorr, JN Rosenau, pp. 87–109. Princeton, NJ: Princeton Univ. Press
- Leyton-Brown D, ed. 1987. *The Utility of International Economic Sanctions*. New York: St. Martin's
- Morgan TC, Schwebach VL. 1995. Economic sanctions as an instrument of foreign policy. *Int. Int.* 21:247–63
- Morgan TC, Schwebach VL. 1997. Fools suffer gladly: the use of economic sanctions in international crises. *Int. Stud. Q.* 41:27–50
- Mueller K. 1998. Strategies of coercion: denial, punishment, and the future of air power. *Secur. Stud.* 7:182–228
- Nossal KR. 1994. *Rain Dancing: Sanctions in Canadian and Australian Foreign Policy*. Toronto: Univ. Toronto Press
- Oppenheim FE. 1981. *Political Concepts: A Reconstruction*. Chicago: Univ. Chicago Press
- Pape RA. 1996. *Bombing to Win: Air Power and Coercion in War*. Ithaca, NY: Cornell Univ. Press

- Pape RA. 1997. Why economic sanctions do not work. *Int. Secur.* 22:90–136
- Pape RA. 1998. Reply to evaluating economic sanctions. *Int. Secur.* 23:195–98
- Preeg EH. 1999. *Feeling Good or Doing Good with Sanctions: Unilateral Economic Sanctions and the U.S. National Interest*. Washington, DC: CSIS
- Schelling TC. 1984. *Choice and Consequence: Perspectives of an Errant Economist*. Cambridge, MA: Harvard Univ. Press
- Schwartz SI, ed. 1998. *Atomic Audit: The Costs and Consequences of U.S. Nuclear Weapons Since 1940*. Washington, DC: Brookings Inst.
- Shultz R, Godson R, Greenwood T, eds. 1993. *Security Studies for the 1990s*. Washington, DC: Brassey's
- Simon HA. 1976. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization*. New York: Free
- Small M, Singer JD. 1982. *Resort to Arms: International and Civil Wars*. London: Sage
- Stam AC. 1996. *Win, Lose, or Draw*. Ann Arbor: Univ. Mich. Press
- Summers HG. 1981. *On Strategy: A Critical Analysis of the Vietnam War*. Washington, DC: GPO
- Tsebelis G. 1990. Are sanctions effective? A game-theoretic analysis. *J. Confl. Resolut.* 34:3–28
- Wang K, Ray JL. 1994. Beginners and winners: the fate of initiators of interstate wars involving great powers since 1495. *Int. Stud. Q.* 38:139–54
- Watson A. 1983. *Diplomacy*. New York: McGraw-Hill
- Wildavsky A. 1979. *Speaking Truth to Power: The Art and Craft of Policy Analysis*. Boston: Little, Brown