Reevaluating Alliance Reliability: Specific Threats, Specific Promises

Brett Ashley Leeds; Andrew G. Long; Sara McLaughlin Mitchell

The Journal of Conflict Resolution, Vol. 44, No. 5 (Oct., 2000), 686-699.

Stable URL:
http://links.jstor.org/sici?sici=0022-0027%28200010%2944%3A5%3C686%3ARARSTS%3E2.0.CO%3B2-0

The Journal of Conflict Resolution is currently published by Sage Publications, Inc..

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/sage.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is an independent not-for-profit organization dedicated to creating and preserving a digital archive of scholarly journals. For more information regarding JSTOR, please contact support@jstor.org.

http://www.jstor.org/
Mon Oct 11 17:29:20 2004

Reevaluating Alliance Reliability

SPECIFIC THREATS, SPECIFIC PROMISES

BRETT ASHLEY LEEDS
ANDREW G. LONG
SARA McLAUGHLIN MITCHELL
Department of Political Science
Florida State University

Previously reported empirical evidence suggests that when conflict arises, military alliances are not reliable; state leaders should only expect their alliance partners to join them in war about 25% of the time. Yet, theoretical arguments explaining the choices of leaders to form cooperative agreements are at odds with such empirical evidence. This puzzling gap between theory and evidence motivates a reconsideration of previous measures of alliance reliability. Many alliance treaties include specific language regarding the circumstances under which the alliance comes into effect, often limiting obligations to disputes with specific target states or in specific geographic areas, and many treaties do not go so far as to require states to join in active fighting. Considering the specific obligations included in alliance agreements provides an improved estimate of the propensity of states to honor their commitments. Results show that alliances are reliable 74.5% of the time.

ALLIANCE FORMATION AND ALLIANCE RELIABILITY

Traditional theories of international relations have stressed the structural impediments to cooperation under anarchy. With no external means of enforcing agreements, leaders find it difficult to make credible commitments to one another. On military security issues, where the costs of honoring commitments are high, reliable agreements should be very difficult to form. Thus, it was not surprising to scholars of international relations to learn that military alliances are reliable only a small fraction of the time. Previous empirical research has suggested that states come to the aid of their alliance partners when war occurs only about 25% of the time (Sabrosky 1980; Siverson and King 1980). This evidence would seem to support the claim that contracts in interna-

AUTHORS' NOTE: This research was supported in part by the Committee on Faculty Research Support at Florida State University. An earlier version of this article was presented at the annual meeting of the Peace Science Society (International), Ann Arbor, Michigan, October 1999. We thank Jeff Ritter for his substantial contribution to the ATOP dataset and comments on this research; Suzanne Werner, Dan Reiter, and Will Moore for helpful comments on previous drafts; Meredith Reid Sarkees and Doug Gibler for consultation regarding alliance data sets; and John Tures for assistance with data collection. Data for replication can be obtained at http://garnet.acns.fsu.edu/~aleeds.

JOURNAL OF CONFLICT RESOLUTION, Vol. 44 No. 5, October 2000 686-699
© 2000 Sage Publications, Inc.

tional politics are worthless and that state leaders view alliances as nothing more than easily dispensable "marriages of convenience."

Yet, if this view is accurate, then one must ask why leaders continue to form alliances. If contracts do not affect behavior in any way, why do state leaders act as if they will? Why do they continue to negotiate agreements and invest great effort in designing precise treaty provisions? Should not the difficulty in enforcing agreements affect the willingness of leaders to form them?

Several scholars have argued that leaders do recognize the difficulty in enforcing agreements under anarchy and choose to form agreements only when they believe there is a reasonable probability that cooperation will be successful (Downs, Rocke, and Barsoom 1996; Fearon 1998; Leeds 1999). In fact, these authors argue that state leaders will take care to design agreements that they think can be fulfilled; agreements that leaders are certain will be unreliable will never be formed. As a result, agreements that are formed are likely to be successful; the fact that cooperating under anarchy is difficult will be reflected in the inability of state leaders to form agreements, not in the failure of the agreements they do form. The inability of leaders to convince potential allies that they will come to their aid if war occurs might well make state leaders reluctant to rely on alliance agreements as a basis for their security policy; fewer alliances are likely to be formed than might be if compliance could be ensured. But leaders who choose to form alliances have already looked ahead and determined that there is a reasonable probability of allied assistance; thus, once alliances are formed, support is likely to be forthcoming. The empirical observation from previous studies that leaders rarely come to the aid of their allies when conflict occurs seems at odds with this proposition.

Even more confusing is why alliance formation would affect the likelihood of war if it were widely understood that alliance agreements are not reliable. Smith (1995) and Morrow (1994) both suggest that the purpose of forming alliances is to signal strength and deter enemies. According to this view, the primary purpose of formalized alliance agreements is to reveal intentions to potential adversaries. But if alliances are usually unreliable, why would adversaries change their behavior as a result of witnessing the formation of an alliance? Morrow and Smith both demonstrate that alliances are credible signals of intention because they do change incentives; forming an alliance makes it more likely that a state will participate in conflict on the side of its ally (see also Fearon 1997). Once again, this seems to suggest that alliance agreements should often be reliable; allies should have incentives to behave as promised.

Smith (1995, 1996) addresses the gap between the empirical observation of many failed alliances and his theory by suggesting that a selection effect produces a biased observed sample. Because reliable alliances are likely to deter conflict, only alliances that are viewed as potentially unreliable by adversaries are likely to be challenged. Because state leaders are more willing to attack states that they think will not receive assistance from third parties, they may target states whose allies they believe will not meet their obligations (see also Gartner and Siverson 1996). Thus, of the alliances that are challenged by adversaries, only a small proportion are reliable, but we cannot generalize from that sample to conclusions regarding the propensity of allies that have not been attacked to fight together.

Although this argument is persuasive, it is not entirely satisfying. If selection effects are at work, the greatest effect should be seen in the choice between targets without allies and targets with allies. If we assume that states only form alliances and expect them to have a deterrent effect if they are usually reliable, then there should not exist a pool of "unreliable" alliances, and it certainly should not be easily identifiable. Given that a contract had been formed and was still in effect, why would an aggressor state assume a contract was invalid? Why should potential adversaries be able to identify alliances that are likely to be fulfilled and those that are not with such success? We argue that the contract itself provides that information to the aggressor. Alliance contracts are specific enough to provide information to a potential attacker about when assistance is likely to be forthcoming and when it is not.

Previous studies have assumed that when states form alliances, they commit to join one another in any conflict that might arise. In fact, most alliance treaties are much more detailed and identify specific conditions under which the alliance comes into effect and specific obligations should those conditions arise. Both allies and attackers know that the commitments among the partners are limited. This may help to explain why attackers seem to be successful at identifying cases in which state leaders will choose not to come to the assistance of their allies. It also suggests, however, that we have been measuring the extent to which states meet their contractual obligations improperly. Only by considering the actual promises made can we evaluate appropriately whether states have fulfilled their commitments.

In this study, we report a new analysis of the reliability of alliance agreements that demonstrates that when circumstances arise that require allies to act, states usually fulfill their promises. The evidence does not suggest that leaders view alliance agreements as mere "scraps of paper" that they can ignore with impunity.[1] It seems more likely that in forming alliance agreements, leaders are careful to specify promises that they expect to be willing and able to uphold. This often means that alliance agreements are not blanket commitments but rather promises of assistance under particular circumstances. It also suggests that alliance agreements may serve to deter only conflicts that meet those specifications.

This article proceeds first with a review and replication of the oft-cited study of alliance reliability conducted by Sabrosky (1980). By using Sabrosky's coding rules, we are able to reach similar conclusions to those that he presented. We then question Sabrosky's operationalization of alliance reliability and suggest an alternative measure that takes into account variance in the content of alliance agreements. We explain the data collection necessary for our analysis and present new empirical results. We find that when war occurs, 74.5% of alliances are reliable. Finally, we discuss the implications of this study for future research.

1. German Chancellor Theobald von Bethmann-Hollweg is reported to have referred to the guarantee of Belgian neutrality as "a scrap of paper" when preparing to invade in August 1914 (e.g., Cooke and Stickney 1931, 382).

DISCUSSION AND REPLICATION OF SABROSKY (1980)

Several early statistical studies attempted to measure alliance reliability (e.g., Singer and Small 1966, 1968; Holsti, Hopmann, and Sullivan 1973). Particularly notable among these is Sabrosky's (1980) finding that allies fight together in war only 27% of the time.[2] Sabrosky argues that state leaders form alliances to accomplish two objectives: to deter war and to win wars should deterrence fail. Sabrosky assumes that regardless of what is stated in an alliance treaty, all alliances are intended to provide states with greater security, and thus the degree to which they have achieved this objective can be measured by whether allies fight together in war. This leads to his choice of research design.

Sabrosky's (1980) unit of analysis is the alliance war performance opportunity. Considering 50 interstate wars and 173 alliances from 1816 to 1965,[3] Sabrosky determines whether any participant in each war was a member of an alliance. Cases in which a war participant belonged to an alliance are considered war performance opportunities and are coded as honored, violated, or abstained. Every war-alliance combination is a distinct case.[4] An alliance is considered honored if at least two of the alliance partners fight together and none fight on the opposite side. An alliance is violated during a war performance opportunity if any signatories fight against one another or if one of the alliance partners is forcibly annexed by its partner. Finally, if no alliance partners participate in the conflict, then the case is coded as abstained (Sabrosky 1980, 168-70).

Table 1 presents our replication of Sabrosky's (1980) reliability results. The first row reports the number of war performance opportunities that involved allies fighting alongside, remaining neutral, or fighting against, as recorded in Sabrosky's Table 6-3. He found that allies honored their agreements (fought alongside) 27% of the time for the entire period from 1816 to 1965, remained neutral 61% of the time, and fought against their allies in 12% of the cases. Our replication focuses on the slightly shorter period from 1816 to 1944.[5] Sabrosky also reports his findings for this time period, and they are quite similar—allies fight together in 28.3% of the cases, fight against one another in 13.1% of the cases, and remain neutral when their partners fight wars 58.6%

2. Siverson and King (1980, 3) report a similar finding for alliance reliability in the same time period (1816-1965); 23% of allies fought alongside each other in war, whereas 77% did not.

3. War data come from Singer and Small (1972), and alliance coding is based on data collected by Singer and Small (1966). Small and Singer (1969) exclude alliances that were formed during or within 3 months prior to the onset of a war. Sabrosky (1980) includes wartime alliances but excludes alliances that were explicit offensive pacts. We do not exclude cases based on either of these criteria.

4. Sabrosky (1980, 167) notes that he adopts the "alliance-as-entity definition employed by Holsti, Hopmann, and Sullivan." Sabrosky does not code whether each alliance participant joined a conflict but rather whether at least one alliance participant joined the fighting. If at least one ally in a multilateral alliance fights alongside a state in a war, the alliance is coded as honored. We also evaluate the performance of an alliance as a whole, but we adopt the convention that if any alliance partner does not fulfill the stated obligations, the alliance is violated.

5. We select this temporal sample because we have coded treaty provisions through 1944. We did, however, replicate Sabrosky's (1980) study with version 2.1 of the Correlates of War (COW) alliance data through 1995; results were similar and are available from the authors on request.

TABLE 1
Replication of Sabrosky (1980)

	Fights Alongside (Honors), n	(%)	*Remains Neutral (Abstains),* n	(%)	*Fights Against (Violates),* n	(%)	*Total* n
Data reported in Sabrosky's Table 6-3, 1816-1965; Singer and Small (1966) alliance data	48	(27)	108	(61)	21	(12)	177
Data reported in Sabrosky's Table 6-3, 1816-1944; Singer and Small (1966) alliance data	43	(28.3)	89	(58.6)	20	(13.1)	152
1816-1944 replication with COW alliance data (version 2.1, June 1996)	51	(27.3)	121	(64.7)	15	(8.0)	187
1816-1944 replication with ATOP data	63	(29.4)	129	(60.3)	22	(10.3)	214

NOTE: ATOP = Alliance Treaty Obligations and Provisions; COW = Correlates of War.

of the time. The results for the period from 1816 to 1944 are cited in the second row of Table 1.

The third row contains our initial efforts to replicate Sabrosky (1980), using the most recently available versions of the Correlates of War (COW) alliance and war data sets. The overall number of war performance opportunities increases from 152 to 187, but this reflects the coding of more wars and more alliances in the updated data. For example, Sabrosky analyzes a total of 50 interstate wars from 1816 to 1965; the updated COW war data contain 52 interstate wars in the smaller period between 1816 and 1944. Regardless, the percentage of allies that honor their commitments is still low (27.3%), whereas the percentage of allies that remain neutral is higher (64.7%), and the percentage of allies that fight against is smaller (8%). The final row presents the replication using the cases included in our newly created ATOP (Alliance Treaty Obligations and Provisions) data set.[6] We find a similar result: 29.4% of allies fight alongside, whereas 60.3% remain neutral and 10.3% fight against.

Sabrosky (1980) discovers that allies fight alongside one another only 27% of the time; the rest of the time, allies do not participate in wars, or they fight against their

6. ATOP stands for Alliance Treaty Obligations and Provisions. It is a data set that includes detailed information regarding the content of alliance treaties (Leeds 2000). We adopt an operational definition of alliances very similar to that established by Singer and Small (1966). An alliance is a written agreement, signed by official representatives of at least two independent states, that includes promises to aid a partner in the event of military conflict, remain neutral in the event of conflict, refrain from military conflict with one another, or consult/cooperate in the event of international crises that create a potential for military conflict. Although our operational definition is nearly identical to that employed by the COW project, 82 cases in the ATOP data set do not appear in the most recently released version of the COW alliance data set (version 2.1, June 1996), and there are 15 cases in the COW data set that we do not include in the ATOP data. We discovered additional cases through historical research that met the established criteria for inclusion and determined after reading primary documents that some of the cases that the COW project had included in its data set did not meet the criteria established by Singer and Small (1966). There are 146 alliances recorded in the COW data set between 1815 and 1944 and 213 in the ATOP data set.

partners. Because Sabrosky considers only joint participation in war to constitute honoring an alliance, he concludes that only 27% of alliances are reliable. Sabrosky's coding rules do not take into account any variation in the content of alliance agreements. Once we consider the specific actions that allies promise to take and the conditions under which the alliances come into effect, the story is likely to appear quite different.

For instance, Sabrosky (1980) includes the second class of Singer and Small (1966) alliances, nonaggression/neutrality pacts, in his analysis. Yet even if the signatories specifically promised to remain neutral in the event of conflict, they are not coded as honoring their commitments unless they fought with their allies in a war. Sabrosky evaluates ententes (consultation pacts) by the same standard, even though consultation pacts are coded as such because the partners did not agree to fight together but only to consult in the event of crisis. Sabrosky (1980, 168) recognizes this problem:

> If we held closely to the formal terms of each alliance, we might well find that most alliance partners could have remained neutral during many (if not all) of their war-performance opportunities without repudiating the specific obligations that they had undertaken to fulfill.

Similarly, Sabrosky (1980) assumes that every alliance requires the partners to fight together regardless of the context of the war. In fact, the conditions under which allies are required to act vary considerably from treaty to treaty; many agreements describe very specific conditions under which the alliance obligations come into effect. Assuming that every alliance applies to every war may result in attributing promises to state leaders that they quite explicitly did not make. Sabrosky does recognize that the specific *casus foederis* of alliance agreements may not have been invoked in the war in question. He argues, however, that regardless of the specific obligations imposed by an alliance agreement, the true goal of alliances is to create security and common defense; thus, it is appropriate to evaluate alliance performance by his standard.

Although it may be the case that state leaders hope that the outcome of any alliance agreement will be assistance in any conflict, and the probability that this goal will be reached is an interesting topic for research, it does not seem appropriate to judge alliance reliability by any standard other than whether the signatories behaved as they promised to behave. In our view, the reliability of agreements can only be fairly evaluated by comparing the obligations they include to actual behavior. Leaders specify obligations carefully for a purpose, and we need to use that information to analyze the effects of these agreements. Our ATOP data set (Leeds 2000) allows us to do that by including more specific information about the obligations of each alliance partner drawn from the content of each military alliance treaty.

RECONSIDERING THE DATA

Our first goal in reevaluating the propensity for state leaders to meet their obligations was to establish what those obligations were. As part of the larger ATOP data collection project, we read 213 alliance treaties signed between 1815 and 1944 and coded

the promises leaders made to one another and the conditions under which these promises were to come into effect.

We started by characterizing the level of assistance promised in each treaty and the conditions under which this assistance was required. The COW alliance coding scheme includes three types of pacts—defense, nonaggression/neutrality, and entente. We found this coding scheme limited because many treaties specify more than one type of obligation. We elected instead to code five dummy variables for each treaty to indicate whether the treaty had any of the following provisions: defense, offense, neutrality, nonaggression, or consultation. This decision allowed for the possibility of coding multiple obligations (e.g., defensive and consultation provisions). Furthermore, we coded whether each of the general obligations was contingent on or limited to eight specific factors: particular adversaries, conflicts in particular locations, a specific ongoing conflict, bilateral or multilateral conflicts, situations in which a potential adversary does not comply with a specific demand, prior agreement for actions taken, an alliance partner being attacked, or an alliance partner being attacked without provocation. We also coded whether obligations varied among alliance members. For instance, some treaties impose defensive obligations on only one member.

We found, in fact, that it is very common for treaties to specify limits to obligations. For instance, a large number of the alliances among European states specify that they apply only to conflicts involving other European powers, not to conflicts outside of Europe. Some alliances specify that defense provisions apply only to aggression by a particular potential adversary (e.g., Germany). Some alliances limit their applicability to a particular conflictual situation and come into effect if an adversary does not meet a particular demand. Neutrality pacts are often contingent on a signatory being the victim of aggression. And treaties often specify asymmetric obligations: for instance, Britain may commit to defend Portugal, Turkey, or Iraq, but the minor powers do not commit to the same obligations to Britain.

Armed with this detailed information with regard to the content of alliance treaties, we returned to the war performance opportunities. From 1816 to 1944, we coded a total of 214 war performance opportunities using 52 interstate wars. Of the 213 alliances in our data set, 137 were in effect when one or more members were involved in a war; this set of alliances had potential war performance opportunities. Yet, we could not assume that the existence of an alliance agreement during a war necessarily meant that a state leader would be violating the alliance commitment were he or she to decline to participate in the fighting. We were careful to consider the conditions under which alliance obligations come into force.

When they sign alliances, leaders specify that the alliance comes into force only if particular conditions are met (the *casus foederis*). If we wish to judge whether allies honor their commitments, we should consider only cases in which the conditions required to activate the alliance are present. When states participate in wars that are outside the domain of their alliance agreements, their allies have made no promises to them, and thus the degree to which allies fulfill promises cannot be judged. We do not consider these cases to represent opportunities to judge war performance and exclude them from the sample. Table 2 shows that in 67 cases, the *casus foederis* specified in the alliance treaty was not met in the war in question.

TABLE 2
Reason Alliance Does Not Require Action in War, ATOP Data, 1816-1944

Specific Casus Foederis *Not Met*	*Nonaggression Pact*	*Consultation Pact*	*Total*
67	14	23	104

NOTE: ATOP = Alliance Treaty Obligations and Provisions.

Describing a case should help to illustrate these circumstances. In 1893, France and Russia formed a mutual defense pact aimed against the members of the Triple Alliance. They specified carefully the actions either would be required to take should the other be attacked by Germany, Italy, and/or Austria-Hungary. During the time this alliance was in effect, Russia fought two Asian wars—the Boxer Rebellion in 1900 and the Russo-Japanese War in 1904. France fought on the same side as Russia in the Boxer Rebellion and did not participate in the Russo-Japanese War. Sabrosky would code this as one honored alliance (for fighting together) and one abstention. We instead consider that the *casus foederis* in this alliance agreement (attack by Germany, Italy, and/or Austria-Hungary) was not met, and thus these two conflicts do not constitute war performance opportunities for this alliance agreement. France and Russia made no promises to one another that were applicable to either of these Asian wars. We exclude both cases from our analysis of alliance reliability.

In addition to considering whether the conditions invoking alliance obligations are met, we also consider the specific actions that states are obligated to take. Some agreements that qualify as alliances include no provisions requiring states to take action in war. For instance, nonaggression pacts simply specify that the alliance partners eschew the use of force in their relations with one another. Every day on which they are not fighting, they are honoring such a promise. Thus, we can code violations of nonaggression pacts any time an ally fights against a partner to whom he or she has promised nonaggression. It would seem odd to code nonaggression pacts as honored if allies fought together or remained neutral. Thus, nonaggression pacts are coded as violated when alliance partners fight one another but are not considered war performance opportunities under other circumstances. Similarly, the reliability of agreements requiring only consultation cannot be judged simply by considering war participation. Thus, pacts that include no promises beyond consultation are also excluded from our war performance opportunities. Table 2 shows that there are 14 opportunities involving nonaggression pacts in which an alliance partner did not fight against its ally and 23 opportunities involving pacts that required only consultation. A total of 104 cases, therefore, do not qualify as opportunities to judge alliance performance in war.

There remain 110 cases that we judge to be instances in which alliance reliability can be appropriately evaluated. In these instances, a war occurs and allies have made promises to act in a particular manner given that eventuality. We consider an alliance commitment honored if (1) an ally fought on the same side with its alliance partner as

TABLE 3
Honored Alliance Commitments, ATOP Data, 1816-1944

Ally Fought with Partner as Promised	*Ally Remained Neutral as Promised*	*Ally Fought with Partner Despite Promising Only Neutrality*	*Total*
43	37	2	82

NOTE: ATOP = Alliance Treaty Obligations and Provisions.

TABLE 4
Violated Alliance Commitments, ATOP Data, 1816-1944

Ally Fought against Partner (Defense/Offense Pact)	*Ally Did Not Aid Partner (Defense/Offense Pact)*	*Ally Fought against Partner (Nonaggression/ Neutrality Pact)*	*Total*
10	7	11	28

NOTE: ATOP = Alliance Treaty Obligations and Provisions.

promised (in a defense and/or offense pact), (2) an ally remained neutral in a conflict as promised, or (3) an ally fought alongside its alliance partner despite promising only neutrality.[7] Table 3 shows that of the 82 cases of honored commitments, 43 involved an ally fighting with a partner as promised, 37 involved an ally remaining neutral as promised, and 2 involved an ally fighting even though it promised only to remain neutral.

We coded an alliance commitment as violated if (1) an ally fought against its partner (in the context of a defense, offense, nonaggression, or neutrality pact), or (2) an ally did not come to a partner's aid even though it had promised such assistance in a defense and/or offense pact. In Table 4, we can see that of the 28 total alliance violations, 10 were cases in which an ally fought against a partner despite having signed a defense and/or offense pact, 11 were cases in which an ally fought against an alliance partner violating a nonaggression or neutrality pact, and 7 were cases in which an ally did not come to a partner's aid after agreeing to in a defense and/or offense pact. It is clear that even with coding rules that take into account the specific promises included in alliance treaties, a significant number of alliance commitments are not fulfilled when war occurs.

Table 5 summarizes our analysis of alliance reliability. Of the 214 total war performance opportunities that are identified according to Sabrosky's (1980) coding rules, 104 (48.6%) either do not apply to the given conflict or do not require action that can be judged by examining war participation. Of those cases for which alliance reliability

7. This last category is one of "super-reliability" and occurs very infrequently in our data (in only 2 of 214 cases).

TABLE 5
Alliance Reliability, ATOP Data, 1816-1944

Alliance Commitment Honored,		*Alliance Commitment Violated,*		*Alliance Does Not Apply*	*Total*
n	*(%)*	n	*(%)*	*(n)*	*(n)*
82	(74.5)	28	(25.5)	104	214

NOTE: ATOP = Alliance Treaty Obligations and Provisions.

can be judged by evaluating war performance, alliance commitments were honored in 82 cases and violated in 28 cases; alliance partners honored their commitments 74.5% of the time. Including alliances that are not relevant to the wars being fought or do not require action in war artificially inflates the total number of war performance opportunities and makes alliance violation appear to be the norm. In fact, when the circumstances specified in alliance treaties arise, most alliance commitments are honored.

In Table 6, we compare the results of our alliance reliability coding to coding using Sabrosky's (1980) rules. Of the 63 cases that Sabrosky would consider honored alliances because alliance partners fight together, we also code 40 as honored. We consider 4 cases that are honored alliances according to Sabrosky's coding rules as violated. These fall into two categories. In 2 cases, we code the alliance as violated because although the partners did fight on the same side in a world war, the ally entered the war years after the partner was attacked and needed assistance. In each of these cases, there is little doubt that the leader whose state was under attack felt that the ally had not met the stated treaty obligations.[8] In 2 other cases, multiparty defense pacts were honored by some partners and not by others. As long as some partners fought together and none fought on the opposite side, the alliance is considered honored according to Sabrosky's rules. We apply a more stringent standard and argue that if any partner did not provide the assistance promised, the alliance was not honored.

We exclude the remaining 19 cases of allies fighting alongside one another from our analysis because the alliance did not require joint participation in the war in question. In 8 cases, the alliance required consultation only; in 3 cases, the treaty was a nonaggression pact; in 3 cases, the partners fought together in an offensive war despite promising support only in the event their partner was attacked; and in 5 cases, at least some partners fought together in a war that did not involve the particular adversary or conflict location that their treaty specified as the casus foederis. These 19 cases represent the circumstance that Sabrosky (1980) hoped would emerge; they are cases in which allies cooperated in war despite the fact that their formal obligations did not require it.

There are 22 alliances that would be coded as violated by Sabrosky's (1980) rules. We also code 19 as violated. The other 3 we exclude from our sample as not

8. These cases are Serbia and Greece in World War I and Poland and Rumania in World War II.

TABLE 6
Comparison of Alliance Reliability Coding: Sabrosky Coding Rules versus Leeds, Long, and Mitchell Coding Rules, ATOP Data (1816-1944)

	Leeds, Long, and Mitchell			
Sabrosky	*Alliance Commitment Honored*	*Alliance Commitment Violated*	*Alliance Does Not Apply*	*Total*
Fights alongside (honors)	40	4	19	63
Fights against (violates)	0	19	3	22
Remains neutral (abstains)	42	5	82	129
Total	82	28	104	214

NOTE: ATOP = Alliance Treaty Obligations and Provisions.

applicable—2 because the treaties obligated the members to consultation only and 1 because the casus foederis was not met.[9]

A very interesting category, however, is the set of cases that would fall into the category of abstentions under Sabrosky's (1980) coding scheme. Most of these are instances in which the alliance treaty does not apply to the conflict in question (82 of 129), which illustrates further the artificial inflation of the total number of real war performance opportunities in the Sabrosky study. In 42 instances, we consider allies to have fulfilled their commitments despite the fact that they were not coded as fighting together in the COW war data. In 35 of these cases, the allies signed a neutrality pact and upheld their agreement by not getting involved. Sabrosky counts only those cases in which allies fight together as honored alliances, which again depresses the percentage of cases in which alliances are honored. We judge remaining neutral when neutrality is specifically promised to constitute fulfilling the alliance commitment. In the remaining 7 cases, the allies participated in the war but did not meet the formal COW criteria for war participation.[10] We code 5 of Sabrosky's abstentions as violated agreements; allies abstained from conflict when they were obligated to participate.

DIRECTIONS FOR FUTURE RESEARCH

Theories developed in the rational-choice tradition assume forward-looking actors. State leaders are expected to consider the anticipated results of their actions in choosing paths to take. It seems puzzling, therefore, first that leaders would commit to agreements only to break those commitments, and second, that if they did so, unreliable agreements would continue to affect international behavior. Why would alliances con-

9. During the war of Italian Unification, France and Austria fought on opposite sides. France and Austria, along with Great Britain, were signatories to an alliance committed to defending the independence and territorial integrity of the Ottoman Empire. Their dispute in Italy was not a violation of the terms of this alliance agreement.

10. A state must sustain a minimum of 100 fatalities or engage a minimum of 1,000 armed personnel in active combat to be considered a war participant by the COW guidelines (Small and Singer 1982).

tinue to be seen as an effective tool of politics if they were so notoriously unreliable? Research on international cooperation and on war behavior suggests that states should often have incentives to fulfill their promises. Yet the previously reported empirical observation that states rarely join their allies in wars was incommensurate with these theories.

By examining past research and the nature of alliance treaties more closely, we determined that the measurement of alliance reliability was flawed. By reconsidering what it means to honor or violate an alliance in the context of what state leaders actually promise, we find that when conditions arise that require allied action, 74.5% of alliance commitments are fulfilled.

Not only does our study help to resolve a nagging puzzle regarding the propensity of states to form agreements that were not fulfilled, but it should also lead to a greater understanding of the correlates of successful and unsuccessful agreements. In this study, we simply identify reliable and unreliable alliances. A clear next step is to explain why some alliances are reliable and others are not. A number of previous studies have attempted with limited success to determine which attributes of alliance participants (e.g., regime type, power status), which characteristics of the international system (e.g., polarity, technological balance), and which properties of alliances (e.g., power relations, size) are related to alliance reliability (e.g., Singer and Small 1966; Russett 1971; Holsti, Hopmann, and Sullivan 1973; Siverson and King 1980; Sabrosky 1980; Krause and Singer 1997; Werner and Lemke 1997; Bolks 1998; Gleditsch 1998). The difficulty in explaining the propensity to honor or violate alliance agreements could be due, at least in part, to the fact that many cases coded as unreliable alliances were not in fact violations of treaty commitments. Improved measurement of the dependent variable may allow scholars to evaluate theories of the causes of compliance and noncompliance with treaty obligations more accurately.

Knowledge that successfully formed agreements are likely to be fulfilled should also encourage scholars to focus more attention on the formation and design of agreements. In light of this reanalysis, it seems more likely that leaders design agreements that they think have a high probability of being reliable. We should not conclude from this analysis that cooperation without enforcement is not difficult. It seems more likely that leaders agree only to cooperative agreements that they expect to be successful. The sample of agreements that state leaders choose to form is biased in favor of those they will fulfill (Downs, Rocke, and Barsoom 1996; Fearon 1998; Leeds 1999). Perhaps the reason that leaders are so careful in specifying the *casus foederis* in their alliance agreements is that they anticipate the development of other conflicts and want to avoid any obligation to participate. Leaders prefer not to make commitments that they will later have incentive to break. This leads to a new focus on which agreements state leaders are willing to form and which ones they are not willing to form.

Our work also has implications for understanding the initiation, escalation, and diffusion of war. For many years, scholars have tried with inconsistent results to determine whether alliances are conducive to war or peace; some find, for instance, that alliances were associated with peace in the 19th century and with war during other centuries (e.g., Singer and Small 1968; Levy 1981). Others have tried to determine which features of alliances make them more likely to have pacific effects and which

features result in conflictual outcomes (e.g., Gibler and Vasquez 1998). Our research suggests that failing to consider the specific conditions to which alliances apply may cause researchers to connect behavior in wars to alliances that are irrelevant to them.

Despite the advances over previous studies that this analysis represents, we recognize that our study has limitations as well. We study whether alliances are reliable only in the context of war performance. In fact, many alliances would require action by alliance partners at levels of force short of war. Examining the role of alliance partners in militarized disputes is one direction for future research. Also, we do not evaluate whether alliance partners have adhered to treaty provisions for low-level actions such as consultation. Some leaders may feel that their allies have violated agreements by failing to coordinate action, but we make no attempt to evaluate adherence to consultation provisions.

Despite these limitations, this study is an advance over previous attempts to measure alliance reliability, and these new measurement techniques do produce significantly different results. Most of the time, state leaders fulfill their alliance commitments. This finding helps us to narrow a troubling gap between theories of international cooperation and past empirical research.

REFERENCES

Bolks, Sean M. 1998. Security policy choices: Foreign policy behavior as a function of threat, capability, and government structure. Ph.D. diss., Rice University.

Cooke, W. Henry, and Edith P. Stickney. 1931. *Readings in European international relations since 1879.* New York: Harper and Brothers.

Downs, George W., David M. Rocke, and Peter N. Barsoom. 1996. Is the good news about compliance good news about cooperation? *International Organization* 50:379-406.

Fearon, James D. 1997. Signaling foreign policy interests: Tying hands versus sinking costs. *Journal of Conflict Resolution* 41:68-90.

———. 1998. Bargaining, enforcement, and international cooperation. *International Organization* 52: 269-305.

Gartner, Scott Sigmund, and Randolph M. Siverson. 1996. War expansion and war outcome. *Journal of Conflict Resolution* 40:4-15.

Gibler, Douglas M., and John A. Vasquez. 1998. Uncovering the dangerous alliances, 1495-1980. *International Studies Quarterly* 42:785-807.

Gleditsch, Kristian S. 1998. Multiple domestic audiences in foreign policy: Democracies, autocracies, and alliance reliability revisited. Paper presented at the annual meeting of the International Studies Association, March, Minneapolis, MN.

Holsti, Ole P., Terrence Hopmann, and John D. Sullivan. 1973. *Unity and disintegration in international alliances.* New York: John Wiley.

Krause, Volker, and J. David Singer. 1997. Patterns of alliance commitments and the risk of armed conflict involvement, 1816-1984. In *Enforcing cooperation: Risky states and intergovernmental management of conflict*, edited by Gerald Schneider and Patricia A. Weitsman, 81-103. New York: St. Martin's.

Leeds, Brett Ashley. 1999. Domestic political institutions, credible commitments, and international relations. *American Journal of Political Science* 43:979-1002.

Leeds, Brett Ashley with Jeffrey M. Ritter, Sara McLaughlin Mitchell, and Andrew G. Long. 2000. *Alliance treaty obligations and provisions (ATOP) codebook.* Unpublished guide, Florida State University.

Levy, Jack S. 1981. Alliance formation and war behavior: An analysis of the great powers, 1495-1975. *Journal of Conflict Resolution* 25:581-613.

Morrow, James D. 1994. Alliances, credibility, and peacetime costs. *Journal of Conflict Resolution* 38:270-97.

Russett, Bruce M. 1971. An empirical typology of international military alliances. *Midwest Journal of Political Science* 15:262-89.

Sabrosky, Alan Ned. 1980. Interstate alliances: Their reliability and the expansion of war. In *The Correlates of War II: Testing some realpolitik models*, edited by J. David Singer, 161-98. New York: Free Press.

Singer, J. David, and Melvin Small. 1966. Formal alliances, 1815-1939: A quantitative description. *Journal of Peace Research* 3:1-31.

———. 1968. Alliance aggregation and the onset of war. In *Quantitative international politics*, edited by J. David Singer, 247-86. New York: Free Press.

———. 1972. *The wages of war, 1816-1965: A statistical handbook*. New York: John Wiley.

Siverson, Randolph M., and Joel King. 1980. Attributes of national alliance membership and war participation, 1815-1965. *American Journal of Political Science* 24:1-15.

Small, Melvin, and J. David Singer. 1969. Formal alliances, 1816-1965: An extension of the basic data. *Journal of Peace Research* 3:257-82.

———. 1982. *Resort to arms*. Beverly Hills, CA: Sage.

Smith, Alastair. 1995. Alliance formation and war. *International Studies Quarterly* 39:405-25.

———. 1996. To intervene or not to intervene: A biased decision. *Journal of Conflict Resolution* 40:15-38.

Werner, Suzanne, and Douglas Lemke. 1997. Opposites do not attract: The impact of domestic institutions, power, and prior commitments on alignment choices. *International Studies Quarterly* 41:529-46.