Analyzing the Dynamics of International Mediation Processes in the Middle East and the former Yugoslavia

Deborah J. Gerner and Philip A. Schrodt

Department of Political Science

University of Kansas

Lawrence, KS 66045 USA

phone: 785-864-3523   fax: 785-864-5700

p-schrodt@ukans.edu   d-gerner@ukans.edu

An electronic version of this paper (PDF format) and the data sets
used in the analysis are available at the KEDS web site:
`http://www.ukans.edu/~keds`

This paper reports on preliminary analyses of our data.
Please contact the authors for updated results before citing or quoting.

Abstract

This paper discusses a new National Science Foundation-funded project that will examine the dynamics of third-party international mediation using statistical time-series analyses of political event data. Third-party mediation was attempted in over half of the conflicts in the post-WWII period and it is likely that the use of mediation has increased following the end of the Cold War. Surprisingly, there have been few systematic studies on mediation. Those that do exist have generally focused on relatively static contextual factors such as the the conflict's attributes and the prior relationship between the mediator and protagonists rather than on dynamic factors—both contextual and process—that may contribute to the success or failure of mediation activities. In contrast, the extensive qualitative literature provides numerous hypotheses about dynamic aspects of mediation. This, however, primarily consists of case studies, often by mediation practitioners, that exhibit little cumulation and, when taken as a whole, are rife with contradictory assertions.

The project will formally test a number of the hypotheses embedded in the theoretical and qualitative literatures on mediation, using automated coding of event data from news-wire sources and employing time-series and event-history methods. A system of specialized event codes that a sensitive to mediation activities will be developed, then events will be coded from news reports using the TABARI machine coding program. The research will look at the factors that influence (1) whether mediation is accepted by the parties in a conflict, (2) whether formal agreements are reached, and (3) whether the agreements actually reduce the level of conflict. The project will initially focus on conflicts in the Middle East, a region where the principal investigators have substantial field experience. After refining the statistical tests on the Middle East case, the analysis will be extended to event data on conflicts in the former Yugoslavia and West Africa.

The paper presents the results of an empirical "plausibility probe" based on existing WEIS-coded event data for the Levant and the former Yugoslavia. It employs a simple measure of third-party mediation efforts as the independent variables and Goldstein-scaled cooperation as the dependent variable. In the Levant, we find a weak but consistent pattern of mediation correlating with past conflictual activity, and resulting in later increases in cooperation. In the former Yugoslavia, the analysis shows strikingly different results for the mediation efforts the UN, European states, and the US. All three respond to increased conflict, but the UN efforts correlate with greater conflict, the US efforts with greater cooperation, and the European efforts have no effect. These results are consistent with many of the qualitative assessments of these efforts, and suggest that the event data approach will produce credible results.

.

This paper discusses a new project that will investigate the dynamics of third-party international mediation in the Middle East, the former Yugoslavia, and West Africa over the past two decades using statistical time-series models of political event data. This project has recently received funding from the United States National Science Foundation and is just getting underway. The paper will focus on the definition of the issues, the underlying theoretical literature, and the anticipated statistical research but includes an empirical "plausibility probe" that uses existing event data.

Following the end of the Cold War the international system has become increasingly vulnerable to sudden outbreaks of serious systematic violence, both intrastate and international. Iraq's invasion of Kuwait, the conflict between Armenia and Azerbaijan, the genocidal violence observed in Bosnia and Rwanda, and the violent internal conflicts in Somalia, Chechnya, Haiti, Algeria, and Liberia are all examples of this. The disappearance of communism as an ideological principle for organizing conflict (and the disappearance of coercive institutions devoted to suppressing ethnic conflict) appears to have "removed the lid" from long-simmering regional and ethnic tensions and stimulated lethal disputes.

One of the most conspicuous responses of the international community to this violence has been mediation. *Mediation is a specific type of political activity that highlights the role of a third party in facilitating a negotiation process*. This "third party" may be a government official whose country is not a direct party to the dispute, someone associated with an international body like the United Nations, or, as in the case of Track II or unofficial diplomacy, representatives of a nongovernmental organization such as the Quakers or an individual such as Jimmy Carter who may or may not be a citizen of one of the conflictual states (Bercovitch & Houston 2000; Irani 1999). Mediation tends to be less formal than arbitration and adjudication. While there are a variety of roles that a mediator may play, s/he generally does not attempt to *impose* a settlement (unlike some other forms of third-party intervention).

Mediation is an extremely common—often but not always successful—form of conflict management. For instance, Butterworth's (1976) study of 310 international conflicts between 1945 and 1974 found that in 82% of the cases there was some form of mediation. In 1983, Kal Holsti reported results of a study that found mediation occurred in 45 percent of 94 post-WWII disputes examined (cited by Bercovitch 1997, 131). In 70% of the disputes that involved mediation, the outcome of the mediation was at least partially "successful." More recently Bercovitch & Houston (1996) identified 241 international disputes in the 1945-1990 period, of which 137 (57%) were mediated at least once.

The literature on mediation generally encompasses three approaches (for general surveys, see Bercovitch & Rubin 1992; Kleibor 1996; Bercovitch & Houston 1996; Starkey, Boyer, &

Wilkenfeld 1999). First, there are many, many case studies (Princen 1992; Zartman 1995; Crocker, Hampson, & Aall 1999; Greenberg, Barton & McGuiness 2000). While some are theoretically-driven (e.g., Wehr & Lederach 1991, Mitchell and Webb 1988), many are fascinating but purely historical narratives. Second, there is a "wisdom literature" on mediation (which typically spans the gamut from negotiating with five-years-olds, through the massive literature on legal, marital and industrial negotiations, and international negotiation) that is also long on antecedents and short on consequences. Fisher's work (Fisher & Ury 1978; Fisher et al. 1997) is probably the best known example of this genre; others include Lall (1985); Brus (1995); Zartman & Rasmussen (1997); and Mitchell (2000). Finally, there is a small quantitative literature that systematically collects data on mediation and tests hypotheses using inferential statistics; this will be discussed below.

Conflicting Claims about Successful Mediation

Finding testable theories in the literature on the correlates of successful mediation can be a bit frustrating. Authors will readily identify what they view to be important variables, but they are much more reluctant to commit to specifying directional relationships. Furthermore, taken as a whole, the field is particularly rife with contradictory assertions:

❖ The best mediators are neutral *vs.* the best mediators have a stake in the issue;

❖ Successful mediation is most likely when power is equal because parties have maximum flexibility *vs.* success is most likely when power is unequal because uncertainty is minimized;

❖ Mediated negotiation works best early in the process, before the parties have taken firm positions *vs.* mediated negotiation works best after the parties have grown tired of the unresolved conflict;

❖ Mediation is more likely to be successful when there is a positive long-term relationship between the mediator and the disputant(s) *vs.* friendship between the mediator and disputant(s) may make mediation more difficult;

❖ A mediator's leverage over the disputant(s) is an important asset to the mediation task *vs.* a mediator's lack of leverage will enhance the prospects for successful mediation;

❖ Active mediation strategies are more effective in international mediation *vs.* mediation will be most successful when the mediator acts as a facilitator and channel of communication rather than using directive strategies.

This ambiguity is not unique to the mediation literature—look at the debate within balance of power theories about whether alliances inhibit or exacerbate conflict or whether extended deterrence actually works—but because of the limited number of systematic studies, the field has not exhibited a great deal of closure.

Quantitative Studies of International Mediation

The quantitative study of international mediation dates back to the 1960s. The initial work was done by Haas (1968, 1986), who focused specifically on the efforts of international organizations to control conflict through mediation and other active measures such as collective security. This work was later extended by Nye (1971) and Butterworth (Haas, Butterworth, & Nye, 1972; Butterworth & Scranton 1980); the Butterworth also included mediation efforts by individual nation-states and by organizations not set up for collective security. Sherman (1987, 1994; Sherman & Neack 1993; Alker & Sherman 1982) further extended this work in the SHERFACS data set. The CASCON data set developed by Bloomfield and his associates (Bloomfield & Leiss 1969; Bloomfield & Moulton 1997) is another resource dating from this period; it shares many of the concepts of the Haas-Butterworth-Sherman effort, notably the coding of "crisis phase" and the categorization of mediator types. Unfortunately, very little statistical work employing contemporary methods has been done with these data collections—Dixon's (1996) study using SHERFACS is one of the few exceptions—and they have largely been used for descriptive rather than inferential purposes.

During the 1990s, the most extensive quantitative analysis of mediation has been in the work of Bercovitch and his associates (Bercovitch, Anagnoson, & Wille 1991; Bercovitch & Wells 1993; Bercovitch 1996a, 1996b; Bercovitch & Houston 1996; Trappl et al. 1997;Wickbolt, Bercovitch & Piramuthu 1999; Bercovitch & Schneider 2000; Bercovitch & Houston 2000). Bercovitch has assembled a data set on mediation efforts for 295 conflicts from 1945 to 1995, and used state-of-the-art statistical methods to test a variety of hypotheses about mediation. This research also demonstrates clearly that there are testable hypotheses in the qualitative literature and identifies many of the key mediation characteristics of theoretical interest.

In addition to these major projects, several other studies have been recently published based on smaller data sets that focus on a limited number of crises. For example, Ayres (1997) looks at the quantitative dynamics of image change in three conflicts; Mooradian & Druckman (1999) do the same for six mediation attempts in the Nagorno-Karabakh dispute. Carment and Rowlands (1998) develop a game theoretic model based on salience, intensity, capability and expected gains for the belligerent, and assess six 1990s crises against the model.

The objective of our proposed research will be to shift from the generally *structural* focus of the Haas-Butterworth-Sherman, CASCON, and Bercovitch studies—which examine the

characteristics of mediators and the conflicting parties—to an emphasis on the *dynamics* of the mediation process as reflected in news reports coded as international event data. In other words, we will be looking at the impact on mediation of variables that change over time. In the qualitative mediation literature, these are generally referred to as "process" variables, although we will also be looking at some dynamic variables that are usually put in the "contextual" category. For example, the relationship between the mediator and a disputant is generally considered a "contextual variable," but it can change at critical moments, as with the December, 1988 decision by the United States to deal directly with the Palestine Liberation Organization (Gerner & Wilbur 2000). We anticipate that this research will fill a gap in the literature between the macro-level variables emphasized in the existing quantitative studies and the micro-level advice to individual negotiators that is found in the "wisdom literature" and the case studies.

Linkage with Other Projects

This proposal grows out of work that was initiated in the summer of 1998 during the "Multiple Paths to Knowledge Project" sponsored by the James A. Baker III Institute for Public Policy, Rice University, and the Program in Foreign Policy Decision Making, Texas A&M University. The overall focus of that project was the challenge of studying international mediation. Charles Taber (SUNY-Stony Brook), Jonathan Wilkenfeld (Maryland), and Schrodt realized that their current research methods—computer simulation using artificial intelligence algorithms, experimental methods, and event data respectively—presented a natural hierarchy of experimental control. Gerner and Schrodt's event data analysis could look at the problem of mediation in a real-world context, but with no controls on the data. Wilkenfeld and his research group (Wilkenfeld et al. 2000) could use a formal experimental design that could control some aspects of the situation (e.g. the tactics used by the mediator and the presentation of the mediation problem) and measure other aspects, such as the "cognitive complexity" of his subjects. But he could not actually find out *why* the subjects were behaving as they did. Finally, Taber's artificial intelligence simulations (Taber 1992, 1999; Taber and Timpone 1994) would give him complete control of every aspect of the situation, including the problem-solving methods used by the simulated actors.

Papers discussing these three approaches were presented at the 1999 International Studies Association meetings in Washington and 2000 ISA meetings in Los Angeles. While we have designed our project so that it can stand alone, in all likelihood we will be comparing our results with these other projects as well as developing tests that can be directly compared with the experimental and simulation approaches.

Event Data and the KEDS Project

This project will use political "event data" to study dynamic mediation processes. Event data—nominal or ordinal codes recording the interactions between international actors as reported in the open press—break down complex political activities into a sequence of basic building blocks (e.g., comments, visits, rewards, protests, demands, threats, military engagements). When aggregated, these data provide summary measures of political activity. Event data will not substitute for a traditional understanding of a political situation, any more than knowing only the price of a company's stock substitutes for knowing about its product, management, and markets. However, while a stock price does not tell *everything* about a company, it tells *something*. If a company is experiencing problems, its stock price will probably fall; if the company's products are in high demand, the stock price will probably rise. The stock prices quoted in daily papers provide summary measures of thousands for companies that an equivalent amount of text could not convey. Event data serve a similar function. The advantages and disadvantages of event data are discussed in greater detail by Azar, Brody & McClelland (1972); Burgess & Lawton (1972); Sigler, Field, & Adelman (1972); Azar & Ben-Dak (1975); and Merritt, Muncaster, & Zinnes (1993).

The KEDS Project

For the past decade, we have been working on the development of the **K**ansas **E**vent **D**ata **S**ystem (KEDS), a computer program that creates event data from machine-readable text (Schrodt, Davis, & Weddle 1994). KEDS is a pattern-matching system using a computational method called "sparse parsing." (Schrodt 2001) Instead of trying to decipher a sentence fully, KEDS determines only the parts required for event coding—for instance, political actors, compound nouns and compound verb phrases, and the references of pronouns—and then employs a large set of verb patterns to determine the appropriate event code. Unlike more complex full parsing, sparse parsing techniques can be used successfully on unedited news wire text. We have experimented with coding a variety of texts, including specialized regional sources in English and German (Gerner et al. 1994). Most of our work, however, has been with Reuters News Service lead sentences. The lead is usually a simple declarative sentence that summarizes the article, e.g., "The United Arab Emirates welcomed a resumption of formal diplomatic ties between Egypt and Syria after a 12-year rift." For closely reported crisis areas such as the Middle East and the Balkans, lead sentence coding provides thorough coverage of political events, but KEDS has also been successfully used to code complete stories in regions that are less well reported, such as West Africa (Huxtable 1997).

Although we have used KEDS to create several data sets, most of our work has dealt with the Middle East. The Middle East exhibits some of the most complicated political behavior in the world, with a variety of state and non-state actors vying for influence in the context of the ongoing

Arab-Israeli conflict and, until 1990, US-Soviet competition. This extensive foreign policy activity presents a realistic challenge to any system of automated coding because of the quantity and variety of material. We have validated KEDS against both the textual record and human-coded events and found no systematic biases in machine coding (Gerner et al. 1994; Schrodt & Gerner 1994). Thomas (1999) found similar results in an independent validation for his KEDS-coded event data on Northern Ireland. When the human and machine-coded data are used in statistical tests, the results are almost indistinguishable except for differences due to the higher number of events in the machine-coded data. An independent test of KEDS by the **P**rotocol for the **A**nalysis of **N**onviolent **D**irect **A**ction (PANDA) project of the Program on Nonviolent Sanctions and Cultural Survival at Harvard's Center for International Affairs found that when coding dictionaries have been optimized for a set of data, the sparse-parsing methods of KEDS can assign event codes to Reuters leads with a 91% accuracy. Typical accuracy is in the range of 80% to 85%, similar to the accuracy of human-coded data (Burgess & Lawton 1972).

In the spring of 2000, Schrodt produced a new automated coding system named TABARI—**T**extual **A**nalysis **B**y **A**ugmented **R**eplacement **I**nstructions—that is based on the same sparse-parsing principles as KEDS (and hence can use dictionaries developed for KEDS) but is far faster and more flexible. KEDS was written in Pascal and worked only on the Macintosh operating system; TABARI is written as "open-source" code in ANSI C++ and is available on the Linux, Macintosh, and Windows operating systems. TABARI eliminates some deep-seated idiosyncrasies of KEDS and is about 70-times faster, reducing the time required to recode a data set from hours to minutes or even seconds.

We originally became involved with machine coding because, after initial start-up costs, it is dramatically faster and less expensive than human coding. Once a researcher has established vocabulary lists of actors and verb phrases, the only significant expense involved in generating event data is the acquisition of machine-readable news reports. Furthermore, a coding system developed at one institution can be used by other researchers through the sharing of vocabulary lists and coding software; this has been part of our collaboration with the PANDA project.

In working with KEDS, we discovered two additional advantages to machine coding. First, it is free of non-reproducible coding biases and is therefore both reliable and transparent. Human coding is subject to systematic biases because of unconscious assumptions made by the coders. For example, Laurance (1990) notes that even expert coders in the military tended to over-estimate the military capability of China in the 1980s because they knew China to be a large Communist country. When event coding is done part-time by students, coder biases are even more unpredictable and difficult to control. In contrast, with machine-coding the words describing an activity will receive the same code irrespective of the actors or time period involved. Any biases

embedded in the machine coding system are preserved explicitly in its vocabulary and can be modified by the researcher; there is no such record in human coding and thus no ability to address this potential problem.

Second, machine coding allows the researcher to experiment with alternative coding rules that reflect a particular theoretical perspective or interest in a specific set of issues. Using contemporary equipment and software (e.g., TABARI running on a 600 Mhz Pentium III), our 100,000-event Arab-Israeli data set can be completely recoded in about thirty seconds. Historically, the most commonly used event data sets for international relations research have been Azar's (1982) Conflict and Peace Data Bank (COPDAB) and McClelland's (1976) World Event Interaction Survey (WEIS). These were both developed during the Cold War and assume a "Westphalian-Clausewitzian" political world view of sovereign states reacting to each other through diplomacy and military threats; they are ill-suited to dealing with ethnic conflict, low-intensity conflict, or multilateral intervention. With machine coding, alternative coding schemes can be implemented and refined with relative ease, as the PANDA project has already demonstrated.

When the KEDS project began in the late 1980s, accurate machine coding was regarded as something that could only be achieved in the distant future. As recently as 1998, an article on early warning dismissed automated coding as something beyond "our current (or foreseeable) knowledge" (Davies & Harff 1998:81).[1] These pessimistic assessments, however, did not take into account "Moore's Law"—the doubling of computer capacity every 18 months—which has made a desktop computer in 2000 roughly 250-times more powerful than a computer in 1988, the year discussions began on NSF's Data Development in International Relations event data project.

With high-capacity computers, automated coding has proven to be an imminently tractable problem. During the past five years, machine-coded data have become completely accepted in the political science community:

❖ Articles using KEDS-coded data have been published in the top peer-reviewed journals in political science, including the *American Political Science Review* (Goldstein & Pevehouse 1997; Edwards & Wood 1998; Wood & Peak 1999; Schrodt & Gerner 2000), *American Journal of Political Science* (Schrodt & Gerner 1994), *Journal of Conflict Resolution* (Bond et al. 1997; Schrodt & Gerner 1997; Pevehouse & Goldstein 1999), and *International Studies Quarterly* (Gerner et al. 1994).

---

[1] Ironically, in December 2000, the University of Maryland's "Global Event Data System"—the project from which this quote derived—shut down its human coding operation and is now moving to the TABARI machine coding system.

❖ At least two dissertations have been completed that develop new event data sets using the KEDS system (Huxtable 1997, regional focus West Africa; Thomas 1999, regional focus Northern Ireland), and we know of additional dissertations in progress at the University of Michigan, The Ohio State University, and Texas A&M University.

❖ Several government projects in the United States and Europe, as well as a United Nations project, have employed machine coding systems (either KEDS, TABARI and the commercial VRA coder now used by the PANDA project) in crisis early warning systems.

In a word, the automated coding of event data has been accepted as both a viable—and in most cases, preferable—alternative to traditional human coding.

New Research

Event Data Coding

As noted above, most work in event data analysis has been done with either the WEIS or COPDAB event coding schemes, which were originally developed in the 1960s. The actors and activities that characterize contemporary crises are frequently quite different than those assumed by these coding systems. For example, almost all post-Cold War crises involve important actors at the sub-national and trans-national level; many involve substantial legitimate or criminal economic activities in addition to military-diplomatic activity. There is a great deal of interest in the event data community in developing new coding schemes to deal with these behaviors.

While we have done all of our earlier work with WEIS, in this project we will work with the new "Integrated Data for Events Analysis" (IDEA) framework under development by Charles Taylor, Craig Jenkins, Joe Bond and Doug Bond (Taylor, Jenkins & Bond 1999; Taylor et al. 1999). IDEA incorporates the event forms of several existing systems—most conspicuously WEIS and *World Handbook of Political and Social Indications* (Taylor & Hudson 1972; Taylor & Jodice 1983)— and provides a cross-mapping back to the original protocols. However, the system also benefits from the experience of 40 years of event coding, incorporates resources such as Princeton's "WordNet" synonym set (`http://www.cogsci.princeton.edu/~wn/`), and is far more flexible and comprehensive that the existing event coding schemes.

We anticipate that if properly constructed and validated, the IDEA system could be a stable system rather than yet-another event coding scheme. Event data research has suffered at least as much from "lock-in" as from scheme proliferation—most international relations data are coded in either WEIS or WEIS-derived schemes, despite the fact that WEIS was developed nearly four decades ago, was intended as merely a first try, and has several well-known weaknesses.

Our hope is that IDEA will reset the baseline and provide a new, and better, standard for future work. (Jenkins, personal correspondence, 20 February 1999)

At present time, the IDEA researchers have focused primarily on the extension of the coding system to handle internal political activities such as protest. While IDEA, like WEIS, contains some categories relevant to mediation, we anticipate that these will also need to be expanded considerably. IDEA uses a multi-level framework that allows events to be coded at increasing levels of detail, and consequently, we can provide customized dictionaries that provide substantial additional information involving mediation activities without requiring modifications in the overall IDEA scheme. We will coordinate with the Taylor et al. group during this process to insure that we are not introducing incompatibilities; Schrodt is already an outside consultant to this project.

In order to use the IDEA coding system, we will need to revise the coding dictionaries for the TABARI program. This involves three steps. First, we will take our existing dictionaries for coding WEIS and make the appropriate changes in the codes: because IDEA is cross-mapped into WEIS, this should be relatively straightforward. Second, we will incorporate any new vocabulary that has been developed by the IDEA project under NSF-funding into the TABARI system. If there are important features of the IDEA system that cannot be readily translated into TABARI, we will add these to TABARI whenever possible within TABARI's sparse-parsing framework. Finally, we will systematically go through news reports from sources such as Reuters and Agence France Presse to determine how third-party mediation activities are described, and add the appropriate vocabulary to our dictionaries.[2]

Variables

Event data provides an extremely rich set of potential variables for the analysis of mediation activities. These include information on the chronology of the conflict, changes in the relations between potential mediators and the protagonists, the initiation and cessation of formal negotiations, and the level of violence between the disputants. Most of the information considered theoretically relevant to the mediation "process" can be coded from event data (as long as the information is reported in news-wire sources), as can quite a few of the "contextual" variables.

In general, our dependent variable will be the success or failure of international mediation. However, as Kleiboer (1996) points out, this can be measured in a variety of different ways. We will look the following measures, among others:

---

[2] At the present time we anticipate producing a single coding system that incorporates all mediation efforts into a single framework. However, another possibility will be to develop two or three different coding schemes that reflect different theoretical perspectives on mediation.

❖ Do the disputants openly agree to mediation?

❖ Do the parties reach formally reach an agreement?

❖ Is the agreement successfully implemented, in the sense that violence is reduced?

These variables capture the main behaviors emphasized in the theoretical literature and can be readily coded using event data. The one potentially important type of behavior that we cannot study is secret negotiations such as those between Israel and the Palestinians that led to the Oslo Accords. Information about such negotiations is usually (if not always reliably) available in the case study literature and we can use that information to cross-check our models. For example, it will be interesting to see if our models predict that something like Oslo should be happening even though the negotiations are not part of the public event record.

As noted above, the literature abounds with suggestions for independent variables. Many of these correspond directly to information coded with traditional event data, including offers of mediation, pressure from allies to support or reject mediation, the promise of economic assistance or imposition of economic sanctions, and the presence or absence of military activity.

As discussed below, we will be using both discrete and scaled event data. In discrete event data, we focus on the occurrence of specific categories of events, e.g. meeting or agreements. Scaled event data, in contrast, aggregate events for a set interval of time (e.g. monthly or biweekly) into a composite measure using a weighted scale such as the one developed by Goldstein (1992) for the WEIS system, or being developed as part of the IDEA system (Taylor et al 1999). Part of our research will also involve experimenting with alternative scales.

Our initial analysis will use event data almost exclusively unless a structural variable is absolutely required to specify a model. Our rationale is three-fold. First, most of the hypotheses that address structural factors are already ably handled by the Bercovitch group; others can be examined through more thorough analysis of data sets such as Butterworth-Scranton, CASCON, or SHERFACS. Second, event data can be inexpensively, consistently, and transparently coded, which allows us to do comparable analyses across multiple cases. (We still have the problem of inconsistent news coverage, particularly in Africa, but this also affects human-coded data.) Finally, the event data are free of the problem of "hind-sight bias" that is an unavoidable risk in human-coded data: Knowing the outcome of a mediation can potentially affect how informed coders assign values to the independent variables. This has not necessarily been a problem in previous research—coding rules and coder training were presumably designed to avoid hind-sight bias, and hind-sight is less likely to be a problem in structural studies than time-series studies—but it would be useful to have a check on this.

That said, we will leave open the possibility of using some structural information when it is clearly dictated by theoretical considerations and can be unambiguously coded. Geographical proximity is one obvious variable; gross indicators of military and economic power (e.g., sufficient to distinguish the U.S. from Kuwait or Nigeria from Liberia) might be another. We will certainly consider the distinction between conflicts that are largely internal (civil wars in Lebanon, Liberia, and Sierra Leone), conflicts involving sovereign states (Israel-Jordan, Iran-Iraq, Senegal-Mauritania), and conflicts involving actors who would like to become sovereign (Palestinians, Kurds, Kosovars). We may be able to accommodate these different types of conflict with a few variables, but it may also be necessary to create separate models. We will try to avoid variables that require judgement calls (e.g. is Russia a "superpower" in the Balkans? Nigeria in West Africa? Does the United States have a "national interest" in Bosnia? Liberia? Israel?)

Regional Focus

Given the amorphous state of theories of international mediation, we intend to first develop our models in regions where we have a substantial knowledge base, then test the more successful of those models in regions where we will depend on other sources for validation.

Middle East

The Middle East is our longest time series (beginning 15 April 1979, a few weeks after the start of the Iran-Iraq War) and it is the region where we have invested the greatest amount of effort in refining our coding dictionaries, often with coders who have had field experience in the Levant. In addition to our quantitative analysis of conflict in the region, Gerner has done field work in the Middle East for close to twenty years and has access to an extensive network of other scholars who do research on the region.

Conveniently for us—if rather inconveniently for the local populations—the Middle East area has experienced a number of conflicts that have been subject to a variety of different mediation efforts and degrees of success. Mediated dyads include:

- Israel and Palestinians
- Israel and Lebanon
- Israel and Syria
- Syria and Lebanon
- Iraq and various international organizations

- Israel and Jordan
- various parties in the Lebanese civil war
- Iran and Iraq
- Iran and the United States

This region has been intensely covered by the international news media and a detailed record of political activity is available. It has also been the subject of numerous case studies of international mediation: An informal survey of the books at the University of Kansas library listed under the

subject heading "Mediation, international" found that fully a third of the case studies dealt with the Middle East.

The Balkans and West Africa

These two contemporary crisis areas have been subject to extensive—but quite different— conflict and mediation efforts. The conflict in the former Yugoslavia was subject to intense international scrutiny from the beginning. Various methods were attempted to mediate or otherwise control instability in the region, including United Nations, European, and U.S. intervention. Like the Middle East, this conflict contains a variety of dyadic disputes that will probably exhibit somewhat different dynamics. The international news media have covered the Balkans extensively, and a number of case studies are available (e.g. Greenberg and McGuiness 2000; Hanson 2000).

The civil conflicts in West Africa, in contrast, have been dealt with primarily through regional intervention by ECOWAS, although more recently there has been some United Nations involvement. We will focus primarily on the civil wars in Liberia and Sierra Leone, although if sufficient data are available, we will also try to look at Senegal-Mauritania, Nigeria-Cameroon, and possibly international efforts to mediate ethnic conflicts within Nigeria.

Unlike the other cases we will study, West Africa is only sporadically covered by the international media (Huxtable & Pevehouse 1996), and case studies of mediation are rare; getting adequate data will be more of a challenge here. Phillip Huxtable's dissertation (Huxtable 1997) included an event-data analysis of regional intervention in West Africa which demonstrated that credible events can be coded for this region, even if the noise level is higher than it is in the Middle East. We already have regionally-specific TABARI coding dictionaries available for these areas.

We have not done fieldwork in the Balkans and have only limited experience in West Africa (albeit we inadvertently found ourselves in the middle of a Senegal-Mauritania border dispute during June 2000). However, the University of Kansas has Department of Education Title VI area studies centers for both Africa and Russia/Eastern Europe and there is substantial area expertise on campus upon which we can draw.

Additional Areas

After we have refined our statistical models on these three regions, we will try to test the models that produced the most coherent results on several other data sets. These additional tests will not be identical to our core tests because of differences in coding systems and the operationalization of some variables, but they will expand the temporal and geographical scope of our analysis. We are currently considering the following sources:

Other Machine-Coded Data Sets

At the moment, KEDS-coded data sets are available on the conflicts between North and South Korea, China-Taiwan, and the civil conflict in Northern Ireland. All of these disputes have involved extensive international mediation. These data sets have been produced through various research projects and dissertation work, and use the WEIS framework or minor variations on WEIS. If the costs of acquiring news-wire text drops substantially—this could occur at any time due to competition in the commercial market for information—we might also code some additional crises on our own: India-Pakistan, Somalia, and the "Great Lakes" region of central Africa (Rwanda, Brundi, Zaire, Congo, etc.) are potential targets.

BCOW

The **B**ehavioral **C**orrelates **o**f **W**ar data set (BCOW; Leng 1987) is a dense, high-quality event data set that focuses on about forty crises over the past two centuries. BCOW uses an extensive set of codes involving mediation activities and comes close to the level of detail found in IDEA. While some of the crises coded in BCOW involve very little third-party mediation, quite a few were mediated—successfully and unsuccessfully—and could be analyzed. BCOW would considerably extend the temporal range of our analysis.

Episodic Data Sets

As noted earlier, a number of data sets are available that focus primarily on the contextual characteristics of international mediation. Unlike BCOW, these do not contain explicit event data, but they do include some information involving the dynamics of mediation. Depending on the models that we find to be most effective in explaining mediation, we may be able to do some ancillary tests on these data sets, which would considerably extend the temporal and geographical scope of our study. The data sets that appear most promising in this regard are CASCON, SHERFACS, and the Bercovitch data set.

The **C**onflict **E**arly **W**arning **S**ystems (CEWS) chronologies (Alker, Gurr, & Rupesingh, forthcoming; `http://www.usc.edu/dept/LAS/ir/cis/cews/`) are another potential source of new data. CEWS provides a series of narratives describing 20 of the major conflicts that were active in the 1990s such as Chechnya, Chiapas, Kashmir, Sudan, and Tibet.

> Designed to illustrate feasibility and usefulness, this prototype reports on 20 conflict cases, using narratives or chronologies provided by an international team of experts in conflict prevention. They were written from a violence diminishing perspective, so that fruitful lessons might be drawn from comparisons of a relatively equal number of relatively successful and unsuccessful cases in the area of the experts' special competence.

(`http://www.usc.edu/dept/LAS/ir/cis/cews/html_pages/aboutcews.htm`;
 accessed 7 August 2000)

While these chronologies are not directly comparable to news-wire based event data—they do not contain multiple reports of events, and generally have a much lower density of events—they still appear to be a potentially valuable resource. In the case of long-running crises such as Kashmir, they contain information going back to the 1940s. The narratives appear quite straightforward to code using TABARI and are already available in machine-readable form on the CEWS web site.

Statistical Models

Much of the prior work in the KEDS project has involved the development (or adaptation) of new computational methods for the analysis of event data. Generally, these methods have come out of the computational pattern recognition literature—for example ID3 (Schrodt 1991a); genetic algorithms (Schrodt 1989); neural networks (Schrodt 1991b); cluster analysis (Schrodt & Gerner 1997; 2000); and hidden Markov models (Schrodt 1999, 2000).

We adopted this approach for several reasons: Pattern recognition was strongly supported by the theoretical literature on political decision-making, many of the pattern recognition algorithms could be employed without the arbitrary intermediate step of scaling the event data into interval-level measures, and with a few exceptions, most of the existing statistical methods used with event data prior to 1990 were very crude, often little more than contingency table analyses. While there has been some additional use of these methods to analyze political behavior—for example neural networks are used by King and Zeng (e.g. Beck, King & Zeng 2000; King and Zeng 2000), genetic algorithms by Sekhon and Mebane (1998) and classification methods by some artificial intelligence researchers (Wickbolt, Bercovitch & Piramuthu 1999; Kovar et al 2000)—computational pattern recognition is still not widely employed in the political science literature. In addition, pattern recognition algorithms lack a clearly defined inferential model, frequently have poorly-understood properties, and, due to the required investment in specialized or custom-written software, the approach is difficult to use without a substantial knowledge of computer programming.

This project, in contrast, will rely primarily on the application of statistical techniques. This is motivated partly by the issue of accessibility, but also by the recognition that the level of sophistication in the time series techniques found in political analysis has increased dramatically in recent years (see, for example, King 1989; Beck & Katz 1995; Box-Steffensmeier & Jones 1997; Beck, Katz & Tucker 1998; Bennett 1999) . Consequently, while time series models are still not a perfect fit to the theoretical explanations for the success and failure of mediation, the inferential power of these methods far outweighs the sacrifices one may need in terms of explanation. Because

these methods can be implemented with existing statistical packages such as *Stata* and *S-Plus*, we will be able to focus most of our efforts on analysis rather than software development.

This emphasis on inferential statistics is also appropriate given the differences between our earlier focus on prediction and the explanatory nature of this project. Good predictive models do not necessarily involve good explanations; in fact when models with diffuse parameter structures are used (for example hidden Markov models, neural networks, or VAR), there may actually be a tradeoff between effective prediction and coherent explanation. Because this project will be evaluating hypotheses form a rich, if inconsistent, theoretical literature, inference is our primary concern.

We are planning to employ a number of different techniques to explore this data. This is *not* intended as a statistical fishing expedition, but rather a recognition that the current state of the theoretical and empirical literature on this topic is not sufficiently focused to dictate a single statistical model. We are well aware that by employing multiple techniques, we are doing multiple statistical comparisons and hence will "beat the significance test" by mere chance on occasion. Our objective is not to find a single best model, but rather to look at the data from a number of different perspectives and, we hope, derive a coherent, statistically-supported explanation of the mediation process.

Interval-level Time Series

A variety of well-understood time series techniques are available for the forecasting of interval-level data: econometricians have been working on these for decades (e.g. Chatfield 1989; Hamilton 1994; Greene 2000). These methods use lagged values of the independent (and some of the dependent) variables to predict current values. In contrast to many economic time series, aggregated event data are very irregular, but this has not prevented the successful application of time-series methods: Goldstein & Freeman (1990); Ward & Rajmaira (1992); Moore (1995); and Goldstein & Pevehouse (1997) are examples of this approach. Because our work is primarily focused on testing hypotheses suggested by the mediation literature, we anticipate specifying our models based on theoretical considerations, rather than using some of the more general techniques (notably VAR) found in earlier event-based studies of international reciprocity.

To use these methods, the events must first be scaled and aggregated. In the work we have done on forecasting, we have obtained credible results using Goldstein's (1992) scale for WEIS data. As noted earlier, the development of scaling is an integral part of the development of the IDEA system, so this scale will be available for the data we code using that system. Following the lead of Goldstein & Pevehouse (1997), we will probably aggregate our data bi-weekly. This gives about 500 data points for the Middle East data, and 250 for the Balkans and West Africa, a sample size

that is sufficiently large that most time series estimation methods can be used with a reasonable level of confidence and the various asymptotic approximations used for the distributions of the estimators should be fairly accurate.

We also will explore the effects of the choice of scales. In a suggestive experiment using alternative scales in a forecasting model employing cluster analysis (Gerner & Schrodt 1998) we found our results to be remarkably insensitive to the scale used to aggregate the data: the simple presence or absence of events provided roughly 50% of the explanatory power of the model. We want to determine whether this results generalizes to other methods.

Time Series Cross-Section Logit

We will use logit analysis to estimate models predicting the success or failure of our three dependent variables (mediation, agreement, and behavioral change resulting from the agreement) as a function of various independent variables. This approach is similar to that employed over the past decade in many of the democratic peace studies: Beck, Katz, & Tucker (1998:1260) list 18 such studies and additional analyses employing time series cross-section logit have appeared more recently. Dixon (1996) applies this method in his examination of mediation using the SHERFACS data. Recent work by Beck, Katz & Tucker (1998), King & Zeng (1999) and others has identified a number of ways to modify the standard logit techniques that can improve the properties of the estimators; we will these suggestions into account in designing our studies.

Event History and Duration Models

The final set of analyses will focus on event history and duration models (see Allison 1984; Blossfeld, Hamerle, & Mayer 1989; Blossfield & Rohwer 1995; Maller & Zhou 1996; Box-Steffensmeier & Jones 1997; Bennett 1999). In these models, the variable of interest is the expected amount of time required for an event to occur, but this is modeled explicitly as a stochastic process rather than as a deterministic process. In other words, the independent variables increase or decrease the probability of an event occurring, but the model does not attempt to predict exactly when the event will occur. This approach is consistent with the theoretical expectations of the mediation literature, which suggests that there is a large random component to the timing of negotiation phases. Duration models have recently been applied by Bennett to study war duration (Bennett 1997; Bennett & Stam 1996) and similar techniques could be used for the study of mediation.

An Initial Test of the Effects of Mediation

In order to demonstrate the credibility of this approach, we have done some initial work on the Levant and former Yugoslavia using existing, WEIS-coded data and a simple definition of mediation. The purpose of this exercise is to provide an empirical "plausibility probe" to demonstrate that non-trivial results can be obtained from event data.

Our dependent variable in this test is the Goldstein-scaled net cooperation scores, aggregated by month, for three time series: Israel's actions towards Palestinians (April 1979 - December 1998), Israel's actions towards Lebanon (April 1979 - December 1998), and Serbian actions towards Bosnia (June 1991- May 1999).[3]   Figures 1, 2 and 3 show these three series.

The independent variables will be the mediation efforts by several different groups. We use a simple—and somewhat indirect—indicator of mediation: *the number of instances where the mediator has a cooperative interaction (WEIS categories 01 through 10) with both sides of the conflict within a period of 7 days*.[4] This pattern does not guarantee that the third party is actually engaged in mediation—and our future work will use more precise measures—but almost all mediation activities will satisfy this criterion. In other words, this measure provides a necessary but not sufficient indicator of mediation activity. The following sets of mediators have been studied:

Isreal-Palestine	USA, UN, EU
Israel-Lebanon	USA, UN, EU
Serbia-Bosnia	USA
Serbia-Bosnia	UN
Serbia-Bosnia	EU, UK, France, Germany, Italy

The analytical technique we use is the cross-correlation between the mediation measure and the level of Goldstein-scaled cooperation in the dyad. In the cross-correlation diagram, the values to the left of zero (the center of the graph) are the correlations with mediation activity and cooperation between the protagonists *prior* to the mediation; the values to the right of zero are the correlations with mediation activity and cooperation *following* the mediation.   If mediation is successful, we would expect to see a positive correlation between mediation events at time t and cooperation at time t+k in these figures. The cross-correlation approach is discussed in greater detail in the Appendix. The cross-correlations were computed using Stata 6.0; the data are available from the authors.

---

[3] The reciprocal measures in each dyad are highly correlated: the two Israel-Palestinian series correlate with r = 0.65, Israel-Lebanon with r=0.70 and Serbia-Bosnia with r=0.58

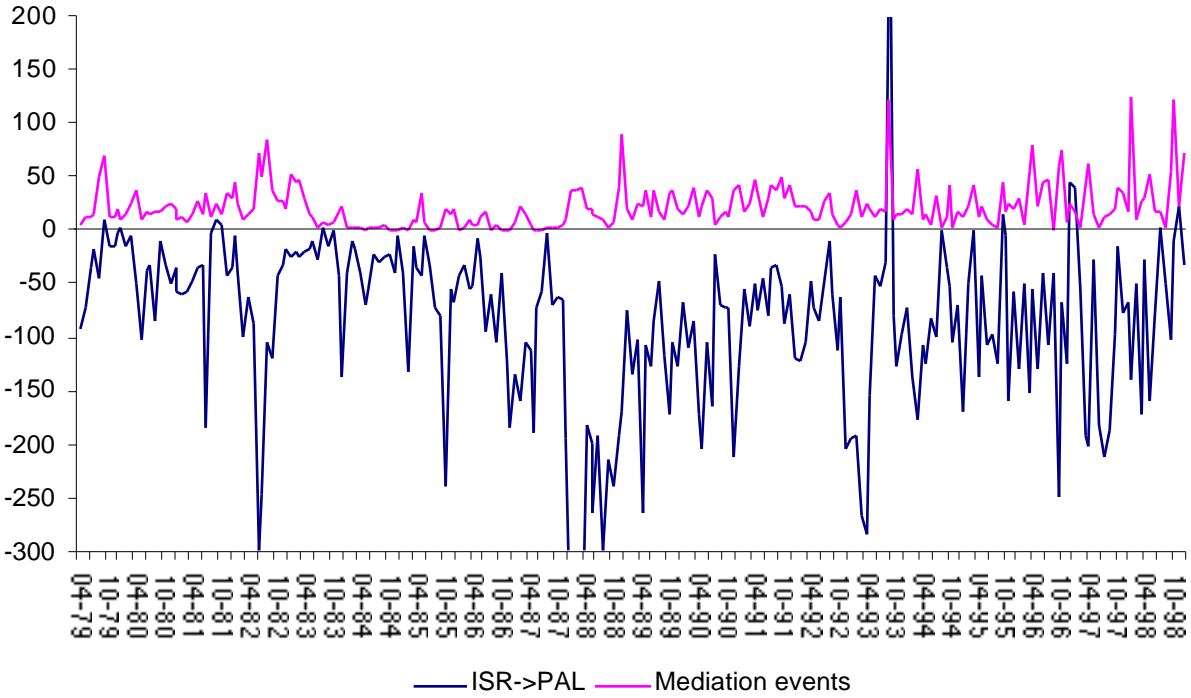[4] We did a few tests using an interval of 4 days; this made no discernible difference in the results.

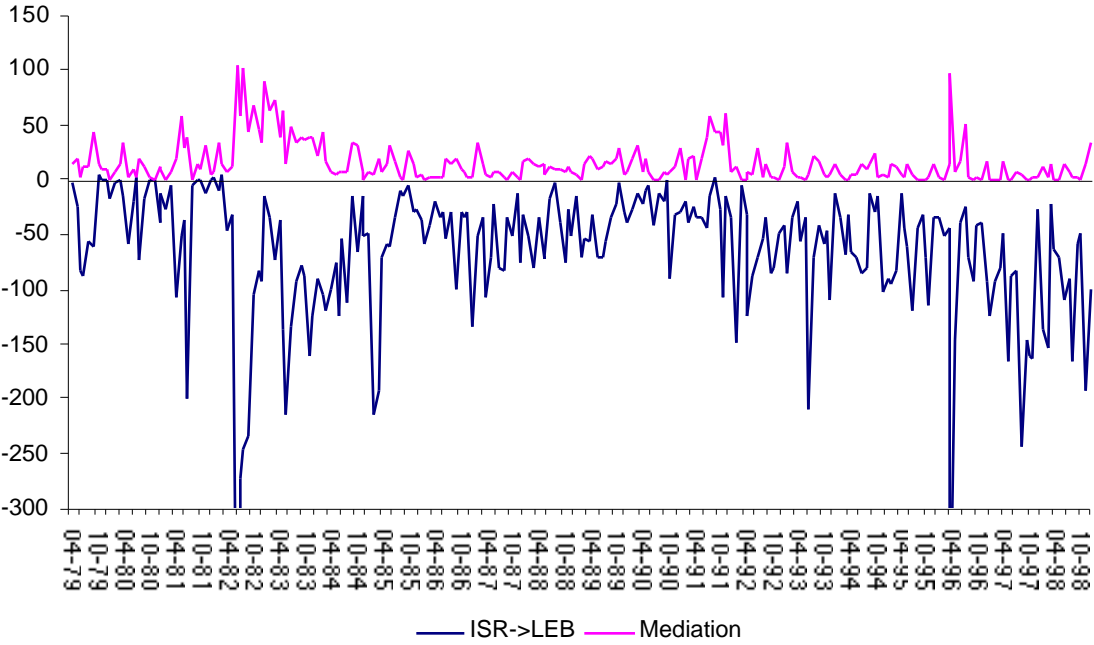**Figure 1. Israel-Palestinian Cooperation and mediation**

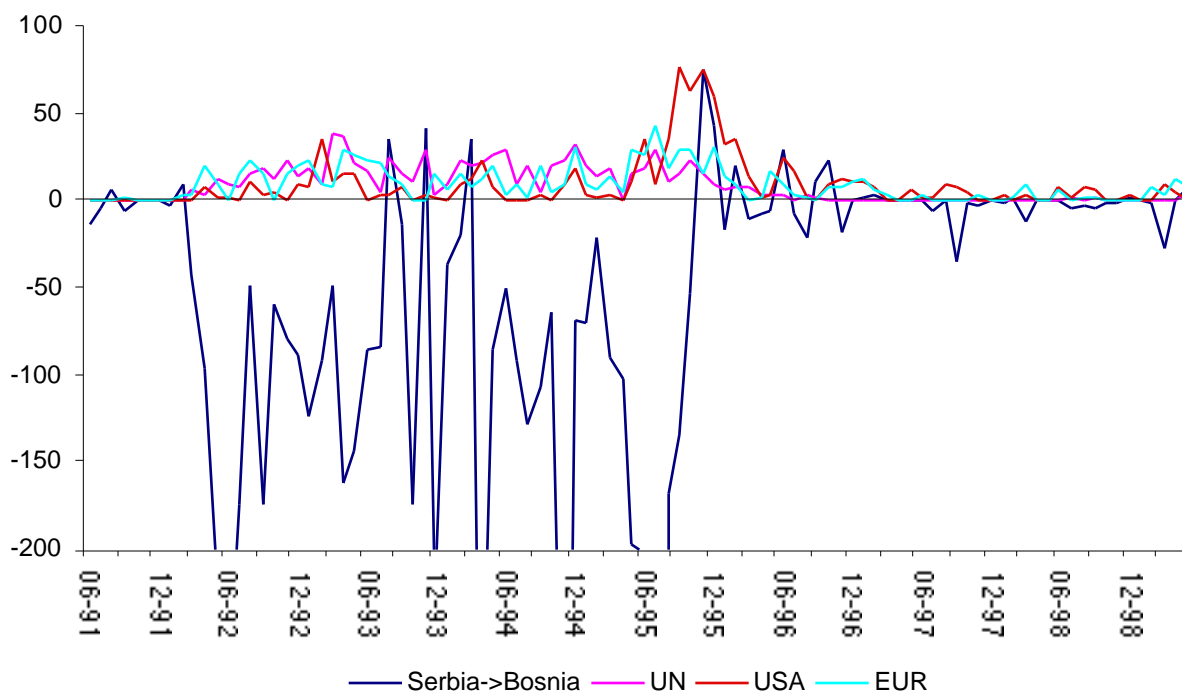**Figure 2. Israel-Lebanon Cooperation and mediation**

Figure 3. Serbia-Bosnia Cooperation and mediation

Levant

In the analysis of the Levant case, we looked at mediation efforts involving either the USA, UN or European Community/Union. Most of this activity, unsurprisingly, involves the USA: of the 95,464 events in the data set, 22752 (23.8%) involve the USA as actor or target; 6186 (6.5%) involve the UN, and only 579 (0.6%) involve the European Community or European Union.[5] Because we looked at interactions involving any of these actors, a meeting between UN officials with Palestinians followed five days later by a meeting between US officials and Israelis would count as a mediation effort. This is imprecise but probably still a reasonable approximation. UN involvement is far more likely in the Israel-Lebanon case than in the Israel-Palestinian case.

Figure 4 shows the cross-correlation of the mediation indicator with Israel-Palestinian conflict measure. The "Mediation" line shows the statistics for one-month measures of mediation; the "5-month MAV" shows the statistics for a five-month centered moving average of mediation (the cooperation variable has not been averaged). The correlogram shows a very distinct pattern of negative correlations for lagged values of cooperation and positive correlations for cooperation in the period following the mediation. In other words, mediation correlates with the level of conflict in

---

[5] In retrospect, there was little point in including the EU as a possible mediator, but the analysis had already been done by the time these aggregate statistics were calculated.

the months before the mediation, and correlates with increased cooperation following the mediation. The levels of correlation are relatively low and most of the positive correlations are probably not individually significant at the 5% level (see Appendix), but the overall pattern is quite regular. The cross-correlations of the centered-moving average is similar to that of the monthly data, but the correlations are somewhat stronger and the pattern is more regular.

As a check that this pattern is actually measuring mediation and not just interactions, we also ran a cross-correlation between Goldstein-scaled cooperation from the USA->Israel and Israel-> Palestinian cooperation. The Goldstein scaled score differs from the mediation score because it only measures interactions between the USA and Israel, without adjusting for whether the US is talking (or otherwise cooperating) with both sides, and also takes into account both positive and negative interactions (e.g. US criticism of Israel). The resulting correlogram—which can be viewed on the KEDS web site (http://www.ukans.edu/~keds/ISA01.supplement/ISA01.Supplement.html) —is quite different than Figure 4: it shows the typical positive spike of contemporaneous correlation at -1, 0 and +1 months[6], but otherwise the correlation is flat and close to zero. We conclude from this that the mediation indicator is picking up something more than simple interaction.

Figure 5 gives the correlogram for mediation and Israel-Lebanon cooperation. This shows a very different pattern that the Israel-Palestinian case. The strongest correlations are contemporaneous—roughly 2 months before and after the mediation—and negative, indicating the mediation is most likely to occur when the level of conflict is high. However, the correlogram gives no evidence that the mediation is effective: the correlations between mediation and subsequent cooperation remain negative, and are generally very close to zero. None of the correlations outside of the contemporaneous period are significant. In short, conflict in the Israel-Lebanon produces mediation efforts, but these have no results.

---

[6] This implies that the US is more likely to direct cooperative behavior to Israel when Israel is cooperating—or at least being less conflictual—towards Palestinians, and more likely to criticize Israel when Israel is being more conflictual towards the Palestinians. While this may seem counter-intuitive given the general perception that the US will support Israel under any circumstances, it is strongly supported by the event data. Furthermore, due to the nature of regression, this says that the US will criticize Israel if Israel actions towards Palestinians are below the mean of those actions. Because that mean is negative (-84.3) and the USR->ISR mean is slightly positive (10.6), the positive correlation means that US criticism only occurs when Israel engages in conflict above a certain level. Another interesting aspect of the data is that the *slope* of the USA->PAL response to PAL->ISR activity is almost twice that of of the USA->ISR response to ISR->PAL activity (0.102 versus 0.058)—in other words, the US reacts more strongly to the Palestinians than to Israel.
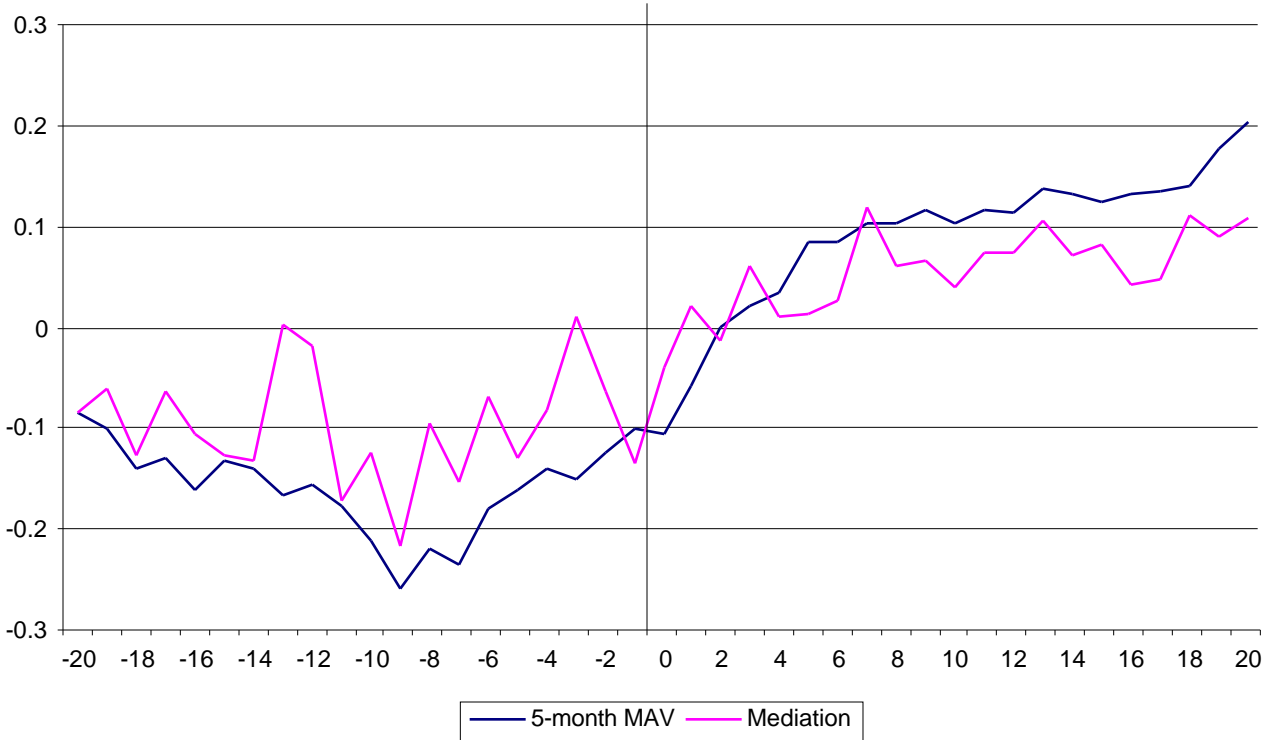
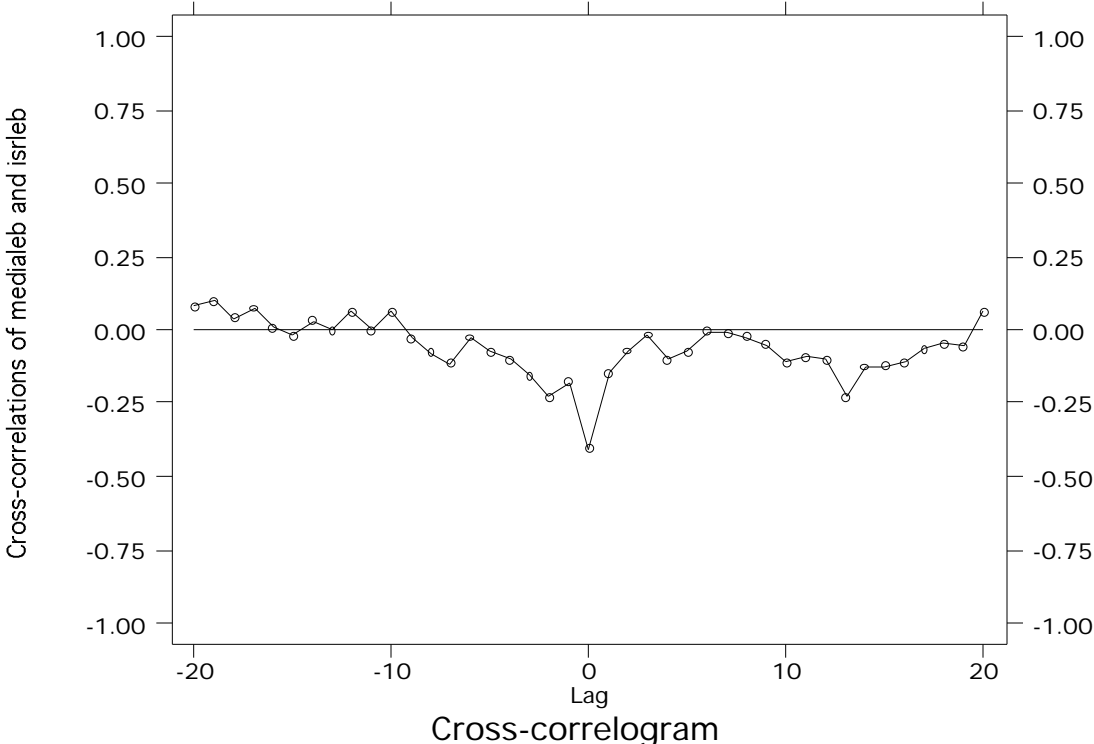**Figure 4. Cross-correlation of mediation with Israel-Palestinian cooperation**

Figure 5. Cross-correlation of mediation with Israel-Lebanon cooperation

We also tested whether Syrian and Iranian mediation had an effect on cooperation from Lebanon to Israel This test was motivated by the suggestion that these states might be effective in reducing conflict between Shi'a groups in southern Lebanon and Israel, and during most of the period being studied, this was the primary source of conflict. The correlogram using LEB->ISR as the dependent variable is basically flat and slightly positive across the full range of lags and leads, indicating a steady level of correlation. However, when ISR->LEB is used at the dependent variable, a pattern very similar to Figure 4 emerges, which again implies weak, but successful, mediation behavior.

This result is somewhat counter-intuitive, and we suspect that it may be due to reporting biases. While the two reciprocal series in the Israel-Lebanon dyad are strongly correlated (r = 0.70) and conflict in this dyad generally follows a tit-for-tat pattern (see Schrodt 1999), the variance of the ISR->LEB series is about seven times higher than that of the LEB->ISR (5574.1 versus 830.0) due to a greater number of news reports involving Israel. As a consequence, the ISR->LEB series may be a more sensitive indicator and therefore show a correlation pattern that the LEB->ISR series does not show, despite the fact that any actual influence Syria and Iran would have in this conflict would be on the Lebanon side.

Former Yugoslavia

Our analysis of the conflict in the former Yugoslavia looked at three different sets of mediators: the United Nations, the United States, and Europe (operationalized as the EU, United Kingdom, France, Germany and Italy).[7] The dependent variable is the Goldstein-scaled cooperation directed from Serbia to Bosnia with the activities of the various ethnic factions combined with those of the governments. In other words, "Serbia" includes the actions of ethnic Serbs in Bosnia and Croatia as well as the actions of the Serbian government; as with all event data, the identification of the ethnicity of individuals or groups responsible for actions was dependent on how the event was reported in the news story. The anecdotal accounts of the conflict suggest that the effectiveness of these efforts varied substantially depending on who was doing the mediation (see Kaldor 1999: 31-68; Weiss 19: 97-136)..

This proposition is supported by our cross-correlation analysis. Figure 6 shows the cross-correlogram of UN mediation and the cooperation measure. The correlations are strongly negative both prior to and following the mediation. In other words, the UN mediation increased following periods of increased conflict in the dyad, but in contrast to the pattern seen for mediation in the Levant, the level of conflict actually become greater following the mediation. US mediation efforts,

in contrast, had a positive effect on cooperation, as shown in Figure 7. As with Figure 6, there is a negative correlation with conflict prior to the mediation, but a strong positive correlation—that is, mediation correlates with increasing cooperation—in the period following the mediation.

Finally, European mediation efforts have no effect, as shown in Figure 8. This shows the usual negative correlation in the lagged period, but contains correlations that are close to zero for activity following the mediation. Unlike the UN efforts, European mediation does no harm, but it does no good either.

Figure 9 looks at another explanation that has been suggested for increased cooperation in the Serbia->Bosnia conflict: the use of force by NATO. "Use of force" was operationalized by events in the WEIS 22 category ("Force") directed from NATO, the USA, UK or France towards Serbia,
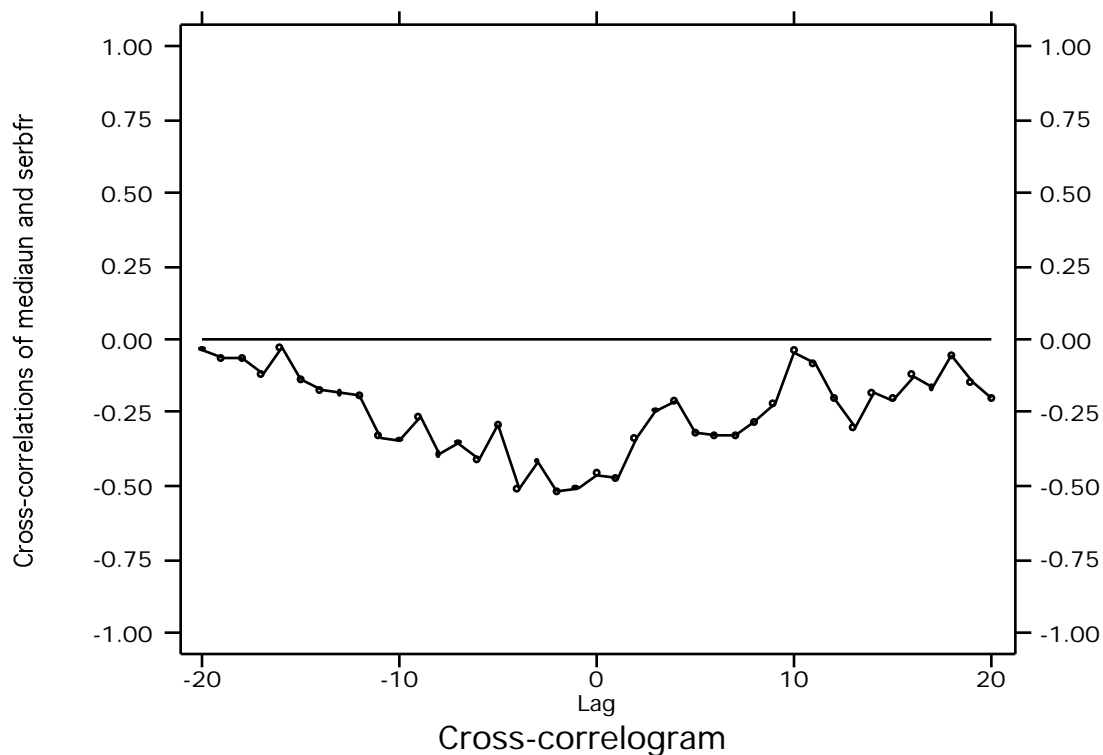

Figure 6. Cross-correlation: UN mediation and Serbia-Bosnia cooperation

---

[7] We also analyzed mediation by Russia and Ukraine. This series has substantially less variance than the European mediation series (17.8 versus 97.6) but generally shows a pattern similar to that of Europe.
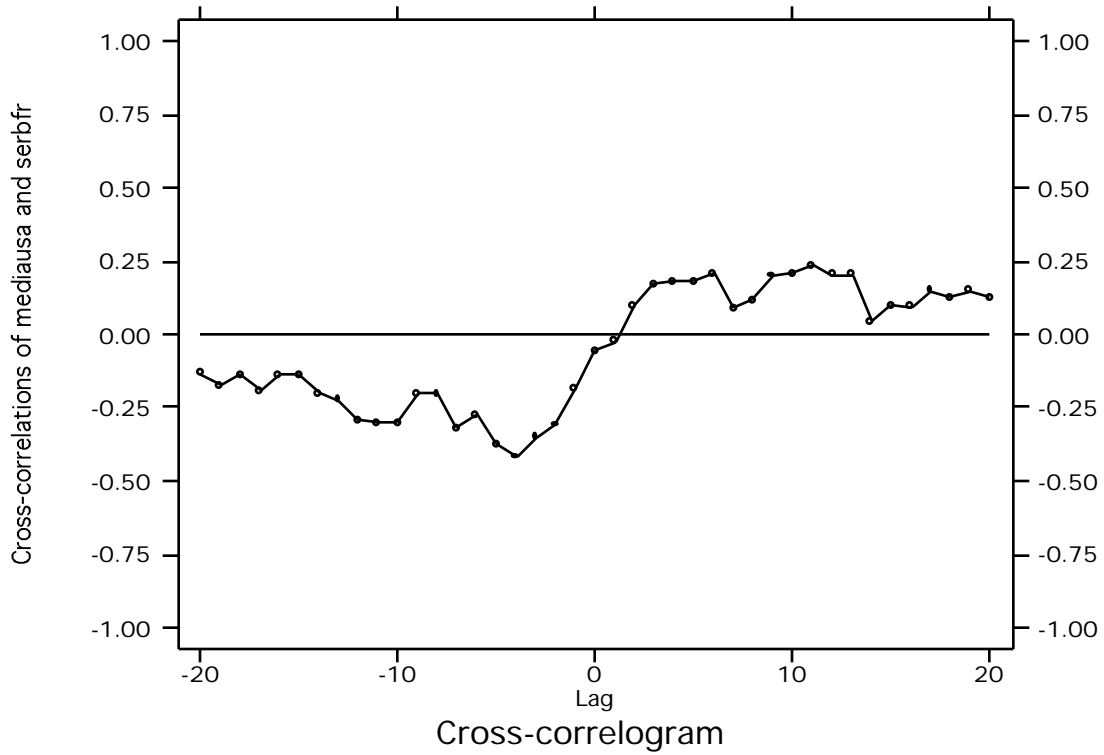
Figure 7. Cross-correlation: United States mediation and Serbia-Bosnia cooperation
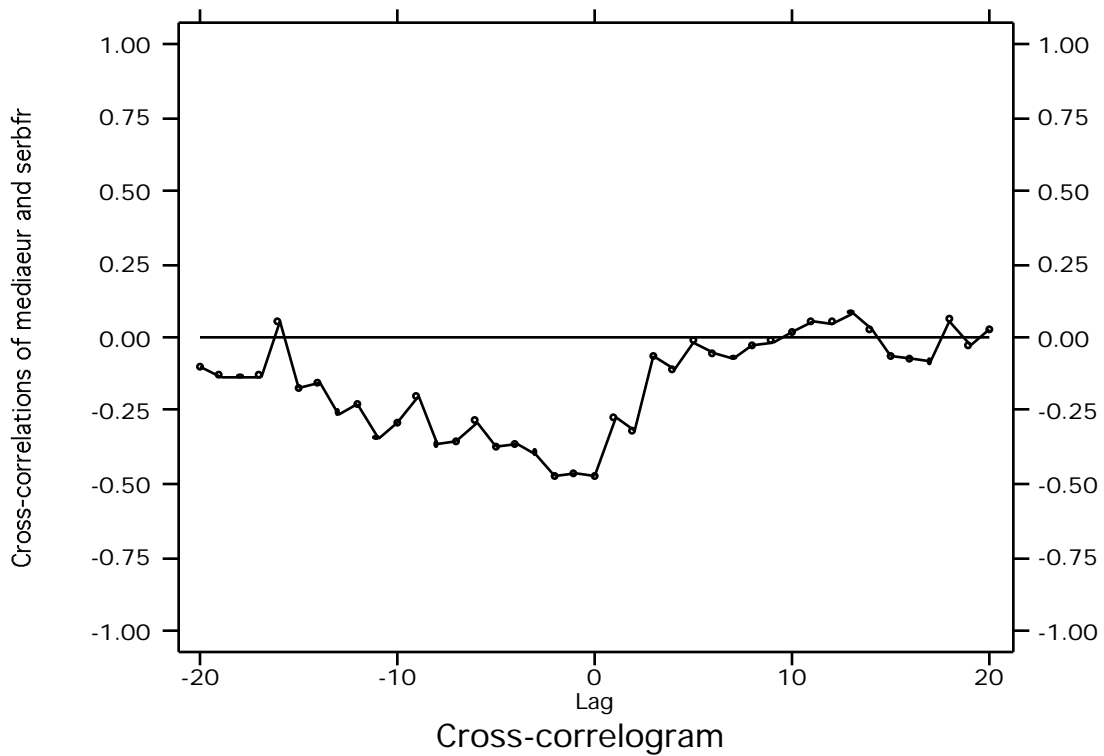

Figure 8. Cross-correlation: European mediation and Serbia-Bosnia cooperation
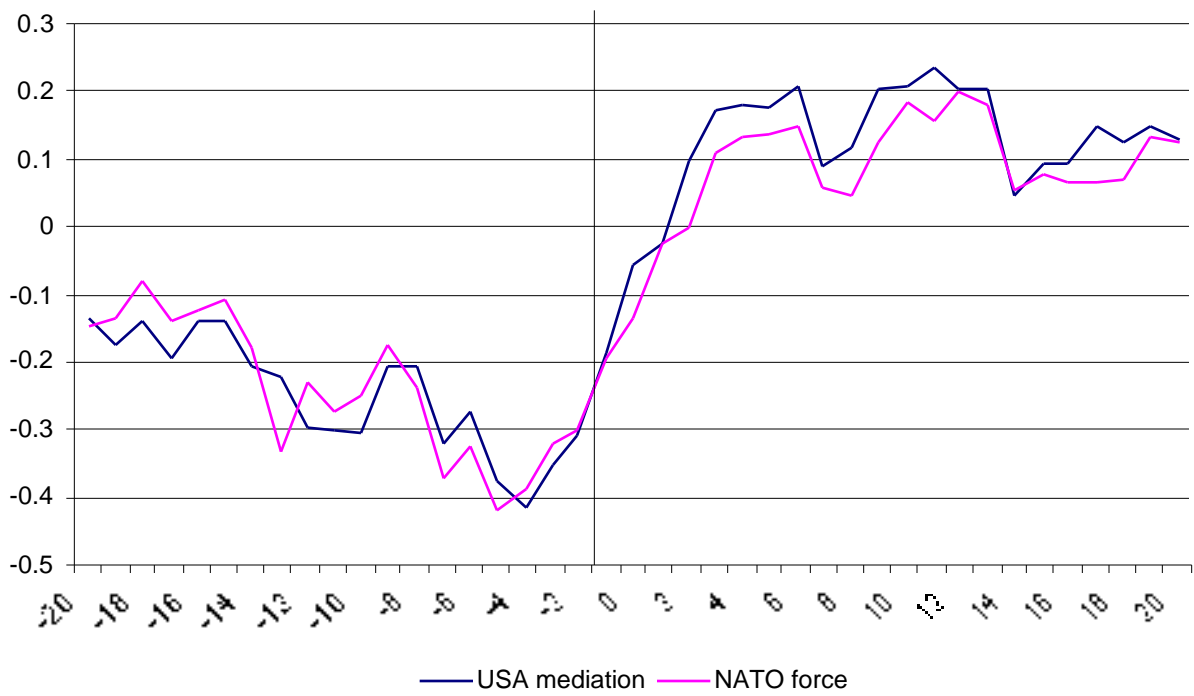
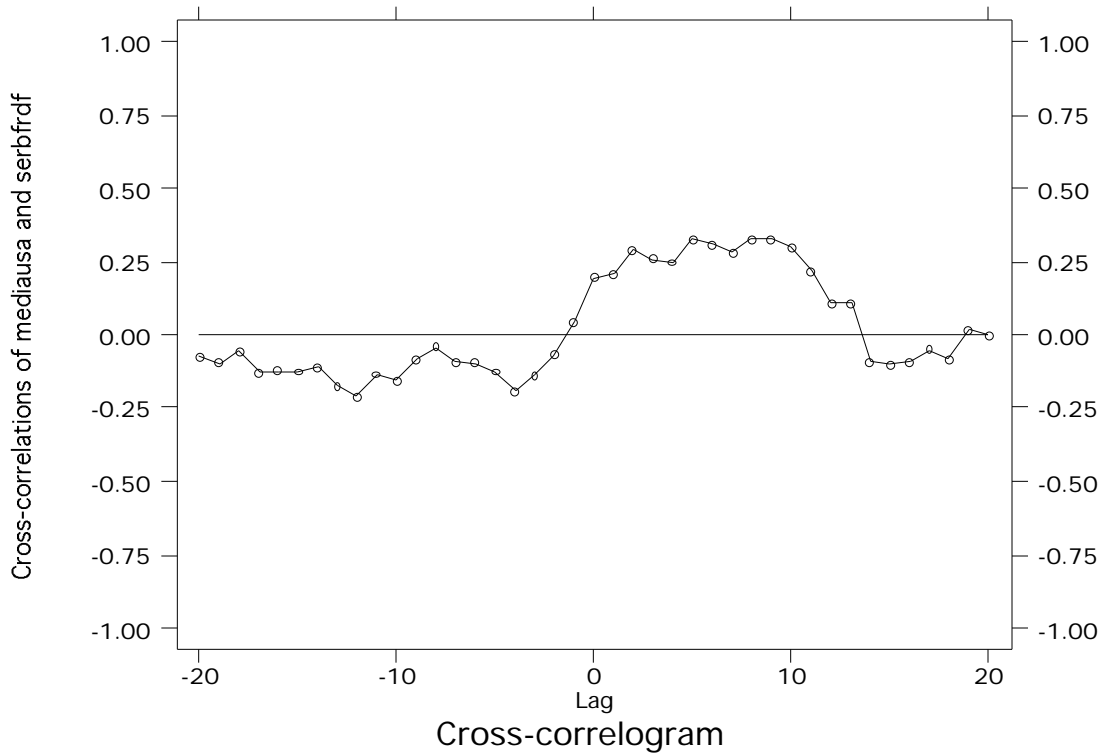**Figure 9. Cross-correlation: Use of force and Serbia-Bosnia cooperation**

Figure 10. Cross-correlation: Use of US mediation and 12-month change in Serbia-Bosnia cooperation

while these actors engaged in cooperative interactions with Bosnia.[8] Figure 9 shows that the two correlograms are similar, but the mediation consistently correlates at a higher level. The "stick and carrot" approach apparently is more effective than the stick alone.

Finally, we looked at the cross-correlation between mediation and the 12-month *change* in cooperation from Serbia to Bosnia; this test is intended to look at the effects of mediation on long-term change rather than the short-term effects on the level of cooperation. Figure 10 is the correlogram for the USA, which shows a consistent level of correlation for a ten month period. The UN and European mediation efforts appear more successful when evaluated by this long-term change measure:[9] The European correlogram for long-term change is similar to that of the correlogram for the United States, though the level of correlation is lower. For the UN, the change in behavior following the mediation has a correlation around zero. This still indicates that UN efforts were ineffective, but by this measure, they at least were not making the situation any worse.

Conclusion

As we noted at the beginning of this paper, the existing theoretical literature on third party mediation contains a wide variety of hypothesis, and we anticipate that a number of additional hypotheses can be derived from the case study literature on mediation. While Bercovitch and his colleagues have done considerable work on structural correlates of mediation success (as did Butterworth and others in an earlier period of empirical work), almost no statistical work has been done on the dynamics of mediation.

Based on the initial empirical work reported here, it appears that event data can be useful for exploring these dynamics. We would emphasize again that we regard the efforts in this paper as very tentative—we used a simple definition of mediation, used only the information available in the WEIS coding scheme, and primarily are using exploratory rather than inferential statistical methods—but the initial results are promising. Distinct patterns appear to be present in the data, and as we refine the coding system to be more sensitive to the nuances of mediation, we expect to be able to go well beyond the statistical description and into hypothesis testing.

---

[8] This is a somewhat quirky measure because it does not include events where force was directed against Serbia but there was no cooperative activity towards Bosnia. The choice of measurement was dictated by the fact that we could calculate this with minor changes in the program used to tabulate mediation; we will look at alternative measures in the future. Given the intensity of international involvement in this crisis, as well as the intensity of news coverage (in the WEIS scheme, a "comment" counts as a cooperative event), it is unlikely that this change will make a great difference in the results, but we will look at it in the future.

[9] Correlograms for the European and UN variables are available on the KEDS web site.

Much of the theoretical literature contains multiple and contradictory hypotheses. From the perspective of understanding mediation, this is a bad thing. From the perspective of statistical research, it is a good thing, for it means that both the validation or falsification of a hypothesis can contribute to supporting some theory. The only "useless" statistical finding in such a situation would be ambiguous or inconsistent results. But our findings thus far are "telling a story", and as we refine both the data and the statistical tests, we expect to see a consistent pattern emerging that should help sort out some of the literature.

Science magazine once surveyed how new techniques in the physical and biological sciences sometimes revolutionized not just the methodologies, but also the theories, within their fields:

> Not everybody appreciates the importance of technique. Many scientists, in fact, are "theory snobs" who dismiss technique as a kind of blue-collar suburb of science. . . . [But there is,] clearly, enormous transforming power in techniques. In the absence of an essential technique, a researcher or a field flounders, developing elegant theories that cannot be decisively accepted or rejected—no matter how many intriguing circumstantial observations are available. But with a key technique in hand, the individual and field move ahead at almost terrifying speed, finding the right conditions to test one hypothesis after another. Conversely, new techniques often uncover new phenomena that demand new theories to explain them. (Hall 1992: 345)

Despite the availability of a large number of descriptive case studies, research on third-party international mediation is arguably theory rich and data poor. Too many theories are chasing too few facts, or at least too little data that have been collected and coded with sufficient consistency that systematic methods can be used to evaluate them. Event data and time series analysis, combined with an appropriate substantive understanding of the political situation that can provide a "reality check" on those statistical results, may be a productive avenue for testing those theories.

27

Appendix: Cross-Correlation

Cross-correlation is not a widely used technique in political science and some explanation is perhaps in order. The measure is similar—but not identical—to computing the Pearson product moment "r" between x_t and $y_{t\pm k}$ for various values of k; both statistics have the form

$$r = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}}$$

The difference between a cross-correlation and a Pearson product moment is that in the cross-correlation Var(x) and Var(y) are estimated from the entire sample, whereas in a Pearson product moment these variances are computed only on the cases that were used to compute the covariance. The technique is useful in determining if a behavior has a long-term effect when the likely timing of that effect is not specified by the theory. Note that the "cross-correlograms" are *not* a time series giving the effect of a single mediation on subsequent behavior; they are a correlation of the mediation with prior and future behavior for the entire time period. For additional information on cross-correlation, see Kendall 1973: 129; Chatfield 1989: 136; and Gottman 1981: 318.

The *approximate* critical value of the cross-correlation coefficient at the 5% two-tailed significance level is $\pm 2/\sqrt{N}$, which is roughly 0.13 for the Levant case and 0.18 for the Balkans case. However, these correlograms have been computed on the raw series rather than the detrended and pre-whitened series (see Chatfield 1989: 137-140) so the correlation may be over-estimated. Consequently, these statistics should be interpreted as primarily descriptive rather than inferential.

We ran cross-correlations on detrended variables, and the results are generally consistent with those found in computations using the raw data; those supplementary correlograms can be found at the KEDS web site. The detrended series, however, still contain autocorrelation at a lags of one month, and sometimes two months, so detrending alone is insufficient to produce a white noise process.

At this point, the cook-book approach would be to continue to process the data until we had "whitened" it. This can be done, but every step in the sequence of standard time-series transformations that improve the statistical characteristics of the estimators—removal of trend, removal of autocorrelation, and the like—also take the data and the analysis further from anything that an analyst can actually understand. For example, when trend and autocorrelation are removed from the time series for US mediation and Serbia-Bosnia conflict (using detrending, then first-difference), then resulting correlogram still has significant negative correlations at lags of -17, -10, -7. -4, -2, -1 and a lead of +1, and positive correlations at leads of +6 and +13. These results are

generally tell the same story as the untransformed data—U.S. mediation responds to past period of high conflict, and has a positive effect on later cooperation.

But the two series on which the statistically correlogram was computed are almost impossible to explain (try it in English…) and one *cannot* say that the correlogram implies that U.S. mediation has a positive effect only at six months and thirteen months. The correlogram implies this is true from detrended and differenced values of that series, a set of transformations that is nearly meaningless from the perspective of figuring out the underlying behavior. The only advantages gained from the transformations are improved analytical properties of the estimators (and even these are just asymptotic approximations). From the perspective of figuring out what was happening in the Balkans during 1990-1999—US mediation improved the situation, UN mediation made it worse—the original data are more useful.

In short, specification of the null model for this data is going to be a complicated process,. While the tendency in statistical analysis in the past has been to pound on the data until it fits some analytical model with known properties (because in the absence of such pounding, one would know nothing about the estimators), contemporary computationally-intensive statistical analysis—for example Monte Carlo and resampling methods—allow one to empirically approximate the properties of estimators based on a much broader set of assumptions. These empirical approximations may, in fact, be more accurate than the analytical approximations, which often are only valid asymptotically. One can do a lot of empirical work with a 1-gigahertz computer, particularly a machine that would otherwise be sitting around searching for alien lifeforms,[10] and and it is likely we will be revisiting this issue in future papers.

---

[10] As in http://setiathome.berkeley.edu

References

Alker, Hayward R., Ted R. Gurr, R.T., Kumar Rupesinghe, eds. forthcoming. *Journeys through Conflict: Narratives and Lessons*. New York: Rowman & Littlefield.

Alker, Hayward R., Jr., and Frank L. Sherman. 1982. "Collective Security-Seeking Practices Since 1945," In *Managing International Crises*, ed. Daniel Frei. Beverly Hills, CA.: Sage.

Allison, Paul D. 1984. *Event History Analysis*. Beverly Hills: Sage

Ayres, William R. 1997. "Mediating International Conflicts: Is Image Change Necessary?" *Journal of Peace Research* 34, 3: 431-447.

Azar, Edward E. 1982. *The Codebook of the Conflict and Peace Data Bank (COPDAB)*. College Park, MD: Center for International Development, University of Maryland.

Azar, Edward E., and Joseph Ben-Dak. 1975. *Theory and Practice of Events Research*. New York: Gordon and Breach.

Azar, Edward E., Richard A. Brody, and Charles A. McClelland, eds. 1972. *International Events Interaction Analysis: Some Research Considerations*. Beverly Hills: Sage Publications.

Beck, Nathaniel, Jonathan N. Katz and Richard Tucker. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42, 4: 1260-88.

Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (and Not To Do) with Time-Series Cross-Section Data." *American Political Science Review* 89, 3: 634-47.

Beck, Nathaniel, Gary King, and Langche Zeng. 2000. "Improving Quantitative Studies of International Conflict: A Conjecture," *American Political Science Review* 94, 1: 21-36.

Bennett, D. Scott. 1997. "Testing Alternative Models of Alliance Duration, 1816-1984." *American Journal of Political Science* 41, 4: 846-878.

Bennett, D. Scott. 1999. "Parametric Models, Duration Dependence, and Time-Varying Data Revisited." *American Journal of Political Science* 43, 1: 256-270.

Bennett, D. Scott, and Allan C. Stam III. 1996. "The Duration of Interstate Wars, 1816-1985." *American Political Science Review* 90, 2: 239-257.

Bercovitch, Jacob. 1996a. "Understanding Mediation's Role in Preventative Diplomacy." *Negotiation Journal* 12, 3: 241-258.

Bercovitch, Jacob, ed. 1996b. *Resolving International Conflict: The Theory and Practice of Mediation.* Boulder, CO: Lynne Rienner Publishers.

Bercovitch, Jacob. 1997. "Mediation in International Conflict." In *Peacemaking in International Conflict: Methods and Techniques*, ed. William I. Zartman, I. and J. Lewis Rasmussen. Pp. 125-153. Washington: U.S. Institute of Peace Press.

Bercovitch, Jacob, J. Theodore Anagnoson, and Donnette L. Wille. 1991. "Some conceptual issues and empirical trends in the study of successful mediation in international relations." *Journal of Peace Research* 28, 1: 7-17.

Bercovitch, Jacob, and Jeffrey Z. Rubin, eds. 1992. *Mediation in international relations : multiple approaches to conflict management*. New York: St. Martin's Press.

Bercovitch, Jacob, and Richard Wells. 1993. "Evaluating mediating strategies: A theoretical and empirical analysis." *Peace and Change* 18, 1: 3-25.

Bercovitch, Jacob and Allison Houston. 1996. "The Study of International Mediation: Theoretical Issues and Empirical Evidence. " In *Resolving International Conflicts: The Theory and Practice of Mediation*, ed. Jacob Bercovitch. Pp. 11-35. Boulder: Lynne Rienner Publishers.

Bercovitch, Jacob and Allison Houston. 2000. "Why Do They Do It Like This? An Analysis of the Factors Influencing Mediation Behavior in International Crises." *Journal of Conflict Resolution* 44, 2: 170-202.

Bercovitch, Jacob, J. and Gerald Schneider. 2000. "Who mediates? The Political Economy of International Conflict Management." *Journal of Peace Research* 37, 2: 145-165.

Bloomfield, Lincoln P., and Amelia C. Leiss. 1969. *Controlling small wars; a strategy for the 1970's*. New York: Knopf.

Bloomfield, Lincold P., and Allan Moulton. 1997. *Managing International Conflict.* New York: St. Martin's Press.

Blossfeld, Hans-Peter, Alfred Hamerle, and Karl Ulrich Mayer. 1989. *Event History Analysis.* Hillsdale, NJ: Lawrence Erlbaum.

Blossfeld, Hans-Peter and Götz Rohwer. 1995. *Techniques of Event History Modeling.* Mahwah, NJ: Lawrence Erlbaum.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor and Kurt Schock. 1997. "Mapping Mass Political Conflict and Civil Society: The Automated Development of Event Data." *Journal of Conflict Resolution* 41, 4: 553-579.

Box-Steffensmeier, Janet M., and Bradford S. Jones. 1997. "Time is of the Essence: Event History Models in Political Science. *American Journal of Political Science* 41, 4: 1414-1461

Brus, Marcel M.T.A. 1995. *Third Party Dispute Settlement in an Interdependent World.* Dordrecht: M. Nijhoff.

Burgess, Philip M., and Raymond W. Lawton. 1972. *Indicators of International Behavior: An Assessment of Events Data Research.* Beverly Hills: Sage Publications.

Butterworth, Robert Lyle, with Margaret E. Scranton. 1976. *Managing Interstate Conflict, 1945-74: Data With Synopses*. Pittsburgh: University Center for International Studies.

Chatfield, C. 1989. *The Analysis of Time Series: An Introduction.* London: Chapman and Hall.

Carment, David and Dane Rowlands. 1998. "Evaluating Third-Party Intervention in Intrastate Conflict." *Journal of Conflict Resolution* 42,5: 572-599.

Crocker, Chester, Fen Osler Hampson, and Pamela Aall, eds. 1999. *Herding Cats: Multiparty Mediation in a Complex World.* Washington: U.S. Institute of Peace Press.

Davies, John, and Barbara Harff. 1998. "Dynamic Data for Conflict Early Warning." In *Risk Assessment and Crisis Early Warning Systems*, ed. John Davies and Ted R. Gurr. Pp. 79-94. New York: Rowman and Littlefield.

Dixon, William J. 1996. "Third-party technique for preventing conflict escalation and promoting peaceful settlement." *International Organization* 50, 4: 653-81.

Edwards, George C., III, and B. Dan Wood. 1999. "Who Influences Whom? The President and the Public Agenda." *American Political Science Review* 93, 2: 327-344.

Fisher, Roger, and William Ury. 1978. *International mediation, a working guide: ideas for the practitioner.* New York: International Peace Academy.

Fisher, Roger, Andrea Kupfer Schneider, Elizabeth Borgwardt, and Brian Ganson. 1997. *Coping with International Conflict: A Systematic Approach to Influence in International Negotiation.* New York: Prentice-Hall.

Gerner, Deborah J., Philip A. Schrodt, Ronald A. Francisco, and Judith L. Weddle. 1994. "The Machine Coding of Events from Regional and International Sources." *International Studies Quarterly* 38, 1: 91-119.

Gerner, Deborah J. and Philip A. Schrodt. 1998. "The Effects of Media Coverage on Crisis Assessment and Early Warning in the Middle East." In *Early Warning and Early Response*, ed. Susanne Schmeidl and Howard Adelman. New York: Columbia University Press-Columbia International Affairs Online.

Gerner, Deborah J., and Ian S. Wilbur. 2000. "Semantics or Substance? Showdown between the United States and the Palestine Liberation Organization." Pew Case Studies in International Affairs, Institute for the Study of Diplomacy, Georgetown University.

Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36, 3: 369-385.

Goldstein, Joshua S., and Jon C. Pevehouse. 1997. "Reciprocity, Bullying and International Cooperation: A Time-Series Analysis of the Bosnia Conflict." *American Political Science Review* 91, 3: 515-530.

Gottman, John M. 1981. *Time-Series Analysis.* Cambridge: Cambridge University Press.

Greene, William H. 2000. *Econometric analysis*. Upper Saddle River, N.J.: Prentice Hall.

Greenberg, Melanie C., and Margaret E. McGuiness. 2000. "From Lisbon to Dayton: International Mediation and the Bosnia Crisis." In *Words over War: Mediation and Arbitration to Prevent Deadly Conflict,* ed. Melanie C. Greenberg, John H. Barton, and Margaret E. McGuiness. Pp. 35-75. New York: Rowman and Littlefield.

Greenberg, Melanie C., John H. Barton, and Margaret E. McGuiness, eds. 2000. *Words over War: Mediation and Arbitration to Prevent Deadly Conflict.* New York: Rowman and Littlefield.

Haas, Ernst B. 1967. *Collective Security and The Future International System.* Denver: University of Denver Monograph Series in World Affairs 5, 1.

Haas, Ernst B. 1986. *Why We Still Need The United Nations.* Berkeley, CA: Institute of International Studies.

Haas, Ernst B, Robert Lyle Butterworth, and Joseph S. Nye. 1972. *Conflict Management by International Organizations.* Morristown, N.J.: General Learning Press.

Hall, S. S. 1992 "How Technique is Changing Science." *Science* 257:344-349.

Hamilton, James D. 1994. *Time Series Analysis.* Princeton: Princeton University Press.

Hanson, Alan. 2000. "Croatia Independence from Yugoslavia, 1991-1992." In *Words over War: Mediation and Arbitration to Prevent Deadly Conflict,* ed. Melanie C. Greenberg, John H. Barton, and Margaret E. McGuiness. Pp. 76-108. New York: Rowman and Littlefield.

Huxtable, Phillip A. 1997. *Uncertainty and Foreign Policy-Making: Conflict and Cooperation in West Africa.* Ph.D. dissertation, University of Kansas.

Huxtable, Phillip A. and Jon C. Pevehouse. 1996. "Potential Validity Problems in Events Data Collection." *International Studies Notes* 21, 2: 8-19.

Irani, George E. 1999. "Islamic Mediation Techniques for Middle East Conflicts." *Middle East Review of International Affairs* 3, 2. http://www.biu.ac.il/SOC/besa/meria

Lall, Arthur S. 1985. *Multilateral Negotiation and Mediation: Instruments and Methods*. New York: Pergamon Press.

Laurance, Edward J. 1990. "Events Data and Policy Analysis." *Policy Sciences* 23:111-132.

Leng, Russell J. 1987. *Behavioral Correlates of War, 1816-1975.* (ICPSR 8606). Ann Arbor: Inter-University Consortium for Political and Social Research.

Mary Kaldor. 1999. *New and Old Wars.* Stanford: Stanford University Press.

Kendall, M. G. 1973. *Time Series.* New York: Hafner.

King, Gary. 1989. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference.* Cambridge: Cambridge University Press.

King, Gary, and Lengche Zeng. 1999. "Logistic Regression in Rare Events Data." Paper presented at the 16th Summer Conference on Political Methodology, College Station, TX, July.

Kleiboer, Marieke. 1996. "Understanding Success and Failure in International Mediation." *Journal of Conflict Resolution* 40, 2: 360-389.

Kovar, K., J. Fürnkranz, J. Petrak, B. Pfahringer, R. Trappl, and G. Widmer. 2000. "Searching for Patterns in Political Event Sequences." *Cybernetics and Systems* 31, 6.

Maller, R. A., and S. Zhou. 1996. *Survival Analysis with Long-Term Survivors*. New York: Wiley.

McClelland, Charles A. 1976. *World Event/Interaction Survey Codebook.* (ICPSR 5211). Ann Arbor: Inter-University Consortium for Political and Social Research.

Merritt, Richard L., Robert G. Muncaster, and Dina A. Zinnes, eds. 1993. *Theory and Management of International Event Data: DDIR Phase II*. Ann Arbor: University of Michigan Press.

Mitchell, Christopher. 2000. *Gestures of conciliation: factors contributing to successful olive branches*. New York: St. Martin's Press.

Mitchell, Christopher and K. Webb, eds. 1988. *New Approaches to International Mediation.* New York: Greenwood Press.

Mooradian, Moorad and Daniel Druckman. 1999. "Hurting Stalemate or Mediation? The Conflict over Nagorno-Karabakh, 1990-1995." *Journal of Peace Research* 36, 6: 709-727.

Moore, Will H. 1995. "Action, Reaction or Rational Expectations? Reciprocity and the Domestic: International Conflict Nexus during the Rhodesia Problem," *Journal of Conflict Resolution* 39, 1: 129-167.

Nye, Joseph S. 1968. *Peace in Parts.* Boston: Little, Brown and Company.

Pevehouse, Jon C., and Joshua S. Goldstein. 1999. "Serbian Compliance or Defiance in Kosovo? Statistical Analysis and Real-Time Predictions," *Journal of Conflict Resolution* 43, 4: 538-546.

Princen, Thomas. 1992. *Intermediaries in international conflict*. Princeton, N.J. : Princeton University Press.

Schrodt, Philip A. 1989. "Short Term Prediction of International Events using a Holland Classifier." *Mathematical and Computer Modeling* 12: 589-600.

Schrodt, Philip A. 1991a. "Classification of Interstate Conflict Outcomes using a Bootstrapped ID3 Algorithm." *Political Analysis* 2: 31-56.

Schrodt, Philip A. 1991b. "Prediction of Interstate Conflict Outcomes Using a Neural Network." *Social Science Computer Review* 9, 4: 359-380.

Schrodt, Philip A. 1999. "Early Warning of Conflict in Southern Lebanon using Hidden Markov Models." In *The Understanding and Management of Global Violence*, ed. Harvey Starr. Pp. 131-162. New York: St. Martin's Press, 1999.

Schrodt, Philip A. 2000. "Pattern Recognition of International Crises using Hidden Markov Models." In *Political Complexity: Nonlinear Models of Politics*, ed. Diana Richards. Pp. 296-328. Ann Arbor: University of Michigan Press.

Schrodt, Philip A. 2001. "Automated Coding of International Event Data using Sparse Parsing Techniques." Paper presented at the International Studies Association, Chicago, February 2001.

Schrodt, Philip A., and Deborah J. Gerner. 1994. "Validity assessment of a machine-coded event data set for the Middle East, 1982-1992." *American Journal of Political Science,* 38: 825-854.

Schrodt, Philip A., and Deborah J. Gerner. 1997. "Empirical Indicators of Crisis Phase in the Middle East, 1982-1995." *Journal of Conflict Resolution* 41, 4: 529-552.

Schrodt, Philip A., and Deborah J. Gerner. 1998. "Cluster Analysis as an Early Warning Technique for the Middle East, 1979-1996." In *Risk Assessment and Crisis Early Warning Systems.* ed. John Davies and Ted R. Gurr. Pp. 95-107. New York: Rowman and Littlefield.

Schrodt, Philip A., and Deborah J. Gerner. forthcoming 2000. "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94,4.

Schrodt, Philip A., Shannon G. Davis, and Judith L. Weddle. 1994. "Political Science: KEDS—A Program for the Machine Coding of Event Data." *Social Science Computer Review* 12, 3: 561-588.

Sekhon, Jasjeet, and Walter R. Mebane, Jr. 1998. "Genetic Optimization Using Derivatives: Theory and Application to Nonlinear Models." *Political Analysis* 7.

Sherman, Frank L. 1994. "SHERFACS: A Cross-Paradigm, Hierarchical and Contextually Sensitive Conflict Management Data Set." *International Interactions* 20, 1-2: 79-100.

Sherman, Frank L. 1987. *Partway to Peace: The United Nations and the Road to Nowhere*. Ph.D. Dissertation, The Pennsylvania State University.

Sherman, Frank L., and Laura Neack. 1993. "Imagining the Possibilities: The Possibilities of Isolating the Genome of International Conflict From the SHERFACS Dataset." In *Theory and Management of International Event Data: DDIR Phase II*, ed. Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes. Pp. 87-112. Ann Arbor: University of Michigan Press.

Sigler, John H., John O. Field, and Murray L. Adelman. 1972. *Applications of Events Data Analysis: Cases, Issues and Programs in International Interaction.* Beverly Hills: Sage.

Starkey, Brigid, Mark Boyer, and Jonathan Wilkenfeld. 1999. *Negotiating a Complex World.* New York: Rowman and Littlefield.

Taber, Charles S. 1992. "POLI: An Expert System Model of U.S. Foreign Policy Belief Systems." *American Political Science Review* 86,4:888-904.

Taber, Charles S. 1999. "Experiments with an Artificial Mediator: A Progress Report." Paper presented at the International Studies Association, Washington.

Taber, Charles S. and Richard Timpone. 1994. "The Policy Arguer: the Architecture of an Expert System." *Social Science Computer Review* 12, 1:1-37.

Taylor, Charles L., and Michael C. Hudson. 1972. *World Handbook of Political and Social Indicators*, 2nd ed. New Haven, CT: Yale University Press.

Taylor, Charles L., and David A. Jodice. 1983. *World Handbook of Political and Social Indicators*, 3rd ed. New Haven, CT: Yale University Press.

Taylor, Charles Lewis, J. Craig Jenkins, and Doug Bond. 1999. "Trends in Protest and Mass Contention: Integrated Data for Events Analysis." Paper presented at the American Political Science Association meetings, Atlanta.

Taylor, Charles Lewis, Joe Bond, Doug Bond, J. Craig Jenkins, and Zeynep Kuzucu. 1999. "Conflict-Cooperation for Interstate and Intrastate Interactions: An Expansion of the Goldstein Scale." Paper presented at the International Studies Association meetings, Washington.

Thomas, G. Dale. 1999. *The "Strange Attractiveness" of Protracted Social Conflict in Northern Ireland.* Ph.D. dissertation, University of South Carolina.

Trappl, Robert, J. Fürnkatz, J. Petrak, and J. Bercovitch. 1997. "Machine Learning and Case-Based Reasoning." In *Learning, Networks and Statistics*, ed. Riccia, G.D., R. Kruse and H-J. Lenz. Vienna: Springer-Verlag.

Ward, Michael and Sheen Rajmaira. 1992. "Reciprocity and Norms in U.S.-Soviet Foreign Policy." *Journal of Conflict Resolution* 36, 2: 342-368.

Wehr, Paul, and John Paul Lederach. 1991. "Mediating Conflict in Central America." *Journal of Peace Research* 28, 1.

Weiss, Thomas G. 1999. *Military–Civilian Interactions.* Lanham, MD: Rowman & Littlefield.

Wickbolt, A., J. Bercovitch and S. Piramuthu. 1999. "Dynamics of International Mediation: Analysis using Machine-Learning Methods." *Conflict Management and Peace Science* 17, 1: 49-68.

Wilkenfeld, Jonathan, Sarit Kraus, Tara E. Santmire and Christopher Frain. 2000. "The Role of Mediation in Conflict Management: Conditions for Successful Resolution." Paper presented at the International Studies Association, Los Angeles.

Wood, B. Dan, and Jeffrey S. Peake. 1998. "The Dynamics of Foreign Policy Agenda Setting." *American Political Science Review* 92, 1: 173-184.

Zartman, I. William, ed. 1995. *Elusive peace : negotiating an end to civil wars.* Washington, D.C.: Brookings Institution.

Zartman, I. William, and J. Lewis Rasmussen, eds. 1997. *Peacemaking in International Conflict: Methods and Techniques.* Washington: U.S. Institute of Peace Press.