

Potentials and Pitfalls in the Application of Event Data to the Study of International Mediation

Philip A. Schrodt
Department of Political Science
University of Kansas
Lawrence, KS 66045 USA
phone: +1.785.864.9024 fax: +1.785.864.5700
p-schrodt@ukans.edu

March 2000

Paper presented at the International Studies Association meetings
Los Angeles, CA

Electronic copies of this paper are available at <http://www.ukans.edu/~keds/papers.html> and
at Columbia International Affairs Online <http://www.ciaonet.org>.

My thanks to Charles Taber, Jonathan Wilkenfeld, Sarit Kraus, Tara Santmire and Christopher Frain for their continued collaboration on this project and to Ze'ev Maoz, Alex Mintz, Cliff Morgan, Glenn Palmer and Richard Stoll for their organization of the "Multiple Paths to Knowledge Project."

ABSTRACT

This paper is a continuation of a joint project involving researchers at the University of Kansas, University of Maryland, and SUNY-Stony Brook that grew out of 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. This paper discusses some issues involved in comparing three methods of studying international mediation: computer simulation, laboratory experiment, and event data analysis. The two major strengths of this comparative approach are (1) substantial similarity between the event-based approaches used in simulation and laboratory experiments, and those used in event data; (2) the large amount of readily-coded information on international mediation available in machine-readable sources. The three major disadvantages are (1) the irregularity of news coverage; (2) the fact that real-world negotiations may be substantially more complex than simulated negotiations and (3) the difficulty in capturing and coding highly detailed aspects of negotiation.

Introduction

This paper grows out of a collaboration that was initiated in the summer of 1999 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 problem of studying international mediation. Charles Taber (SUNY Stony Brook), Jonathan Wilkenfeld (Maryland), and I decided that our current research methods—computer simulation using artificial intelligence algorithms, experimental methods, and event data respectively—presented a natural hierarchy of experimental control. By using event data I could look at the problem of mediation in a real-world context. With that approach, I had no controls on the data, and simply had to take what the historical record gave me. Wilkenfeld could use a formal experimental design, where he could control some aspects of the situation (notably 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 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 ISA meetings in Washington. In brief discussions following those presentations, and in some extensive email correspondence subsequent to the ISA meetings, we found that what appeared at first to be very comparable approaches were in fact more dissimilar than we had anticipated. Nonetheless, the [naïve?] elegance of our original design continues to attract us to the possibility of actually instituting this comparative study. While our 2000 ISA panel was originally intended to present initial results, we have instead transformed it into a round-table discussion of the methodological issues involved in implementing a comparison across these three very different methods of analysis.

This paper will address the "potentials and pitfalls" involved in applying event data to the study of international mediation. Some of my remarks will deal with problems with event data in general—for example the irregularity of news coverage—and some will involve specific comparisons with the Wilkenfeld et al (2000) research design. Comments on the simulation method will be based

on Taber 1999, as well as Taber's earlier work on the POLI artificial intelligence model (Taber 1992; Taber and Timpone, 1994). Most of the comparisons with the experimental approach also apply to the simulation; I refer to these jointly as "laboratory" techniques.

I am focusing my substantive discussion primarily on the two empirical cases that I am hoping to study, the Arab-Israeli dispute and the Balkans conflict. As our joint discussions proceed, it may turn out that these two cases are not good choices for comparison. Both are somewhat atypical in that they have had substantial major-power involvement, and they are highly militarized, protracted territorial disputes. Both of these characteristics make these cases quite distinct from the fisheries mediation—modeled on a dispute between Canada and Spain—that is being studied by the Maryland group, though less so from some of the cases analyzed in Taber's POLI experiments, which looked at US policy towards Asia in the 1950s. However, the event record for these negotiations is fairly complete because they have been densely covered by the international media, and I already have the data.¹

I should note from the outset that while there may appear to be formidable obstacles to making comparisons between these levels of research, I am still optimistic about this project as a whole. It is likely to be hard, but not impossible. Furthermore, as I will note in the conclusion, trends in natural language processing and the availability of machine-readable reports of political activity may make it even more feasible in the future. The very fact that we are even thinking about an exercise such as this may indicate some maturity in the discipline of the scientific study of international relations.

The Experimental World versus the Real World

The experimental world and the world outside of the laboratory—here referred to as the "real world"²—are quite different. This is not unique to the issue of international mediation, or even to

¹ These data sets are available on the Web at <http://www.ukans.edu/~keds/data.html>.

² I am taking the rather unfashionable materialist position that there exists an objective reality that is independent of any individual who observes that reality. Human behavior included. This is *not* the same as saying that I have *unbiased access* to that reality, nor that I think that Reuters reporters or *New York Times* editors have such access. They, as I, filter their perceptions of that reality through preconceptions, cultural assumptions, subconscious interpretations and so forth, and their textual reports will reflect that. I was taught this in sixth grade. But the objective world is out there nonetheless. Those who find this proposition controversial might be advised to spend less time in libraries, seminar rooms and international conferences held in five-star hotels, and more time in the field directly observing human political behavior.

the social sciences. Generalizing experimental and simulation results to messier and less controlled environments has been a problem ever since these methods of scientific inquiry were developed.

Nonetheless, if one wishes to assess the extent to which the experimental and simulation results to correspond in some way with actual human political behavior—and this is clearly our mutual objective—it is necessary to look at comparisons between the two domains. In this instance, I am specifically interested in the comparison of the laboratory results with the real world as that world is viewed from the perspective of the international news media whose reports are the basis of event data.

Event data analysis is not the only way that the real world could be compared with laboratory results. For example an equally legitimate (and quite possibly more valid) approach would be to look at comparisons with the transcripts of actual negotiations where these are available—this approach was employed in the studies in Axelrod 1976. In some cases, a researcher might be able to gain access to negotiations and code those interactions from direct observation. But the present project is using event data as the indicator of real world behavior.

As will be evident from the discussion below, news reports and laboratory data are quite different, but not on a single dimension. In particular, one is not uniformly “simpler” than the other. News reports contain less detail on the meticulous proposals and counterproposals of negotiation, but—reflecting the complexity of the empirical world—these reports are likely to be much more complicated in terms of the number of issues that are dealt with, and quite possibly in the negotiation strategies employed. Laboratory data, in contrast, provide almost unlimited levels of detail in the negotiating process (in theory if not always in practice) but it is very difficult—and possibly unwise—to move the complexity of the experimental or simulated process to the level of that would be found in an actual international negotiation.

Potentials

1. Machine coding from machine readable natural language sources

The bulk of human interactions utilize, and are recorded in, natural language. But to date most systematic studies of human behavior that follow the scientific model of research have analyzed numbers. In some fields of study—most notably economics and geography—this disjuncture is only mildly inconvenient, for most of the behaviors of interest can be represented numerically with little distortion. But in other fields—for example, anthropology,

communications studies, psychology, political science, and sociology—considerable information is lost in the transition from a natural language representation of behavior to a numerical representation. Furthermore, because those transitions are usually mediated through human coders, who interpret the information through a haze of preconceptions, misunderstandings, personal biases, and sloth, substantial error is introduced into the data.

The World Wide Web and the personal computer revolution have made available in machine-readable form a vast amount of information about social behavior instantiated as natural language texts. This information ranges from routine news reports to political speeches, debates, and policy statements to the manifestos of millenarian UFO cults. In addition to Web-based sources, many other materials such as transcripts, interviews, protocols, and treaties are either available in electronic form or readily converted to that form.

Until recently, social scientists had only a limited number of tools to systematically locate and analyze this information; this stands in contrast to the massive efforts devoted to developing software for the analysis of numerical information. Computational tools employing simple pattern recognition—for example, keyword-in-context (KWIC) indices, Boolean search engines and a variety of systems for tabulating words and phrases—were available from almost the beginning of computational social science, but few of these systems made even the most basic linguistic distinctions such as differentiating between subjects and objects in sentences. In retrospect, the computational power required for the sophisticated analysis of textual material simply was not available.

During the 1980s and 1990s, there was an exponential expansion in the capacity of personal computers. This was accompanied by substantial work in computer science and linguists on natural language processing (NLP), and computer programs are now capable of performing many tasks once assumed to require human attention. Grammar checking and document summary, for example, is now a routine feature of many word-processing programs; automated indexing and document retrieval are widely used in specialized applications. Inexpensive language translation programs—a task abandoned as impossible in the 1950s and 1960s—now handle translations of technical documentation, web pages and business correspondence. High capacity computers can accommodate the large numbers of grammatical rules, idiomatic expressions, and context-sensitive substitutions that are necessary to produce credible translations where earlier systems, constrained to simpler rules and word-for-word dictionary translations, had failed.

None of these systems are perfect. Machine translations are usually inferior to those of a fluent translator, and break down completely when dealing with oddly constructed or excessively complex sentences as might be found in poetry, classical literature or political rhetoric. But machine translation is typically superior to that provided by a translator with limited fluency, and when translating routine material such as maintenance manuals, a customized translation program with a broad range of specialized vocabulary may actually produce a document superior to that provided by a human without the domain-specific knowledge.

This change in the ability of machines to handle NLP problems was reflected in event data research in political science: In 1990 almost all event data projects used human coders, whereas in 2000 almost all projects used automated coding.³ Projects that once would have required tens of thousands of dollars, a flock of student coders with a complex supervisory infrastructure and months of painstaking effort can now be done by a single researcher in a few days or weeks, provided appropriate pre-existing dictionaries were available.⁴ This development has, in turn, led to a series of less-obvious spin-offs such as the ability to easily experiment with new event coding systems such as the Protocol for the Analysis for Nonviolent Direct Action (PANDA) and with the coding of internal actors (Gerner et al 1994, Schrodt and Gerner 1994, Huxtable 1997, Goldstein and Pevehouse 1997, Thomas 1999).

Improvements in communication and computer technologies also have dramatically changed the quantity and timeliness of the information available for use in event data analysis and early warning. Material relevant to political behavior is available from the commercial efforts of Reuters, *Agence France Press* (AFP), and other news agencies, and from the equally vast, if more specialized, networks of intergovernmental and nongovernmental organization (IGO and NGO) fieldworkers such as the UN Department of Humanitarian Affairs "ReliefWeb" site (<http://www.reliefweb.int>).

The question remains as to whether these sources are sufficient to understand international events. Whaley notes:

³ This is particularly noticeable in the research published in top journals such as the *American Political Science Review*, *American Journal of Political Science* and *Journal of Conflict Resolution*. Human-coded data has continued to be employed in some government-sponsored projects, most notably the State Failures Project (Esty et al 1995, 1998), but even these efforts are gradually moving to automated coding.

⁴ Huxtable 1997 and Thomas 1999 are examples of dissertations based on large-scale event data sets generated by a single researcher using fully automated methods; Reising 1999 used machine-assisted methods but these were still substantially faster than traditional methods.

The international and major independent news media are, in a practical sense, intelligence services, and they are so used by most persons regardless of their station in the polity or their access to conventional intelligence sources. Indeed, the "prestige" newspapers are known to draw readers from a far higher proportion of senior government officials than of the literate adult citizens in general...

International news media resemble professional intelligence services in both function and structure. They ferret out, collect, collate, evaluate, analyze, summarize and report vast quantities of information, and they do so with an organization comprising correspondents in the field with their local networks and stringers. (Whaley 1973, 238-239)

Political and technological changes since Whaley made this statement—the end of the Cold War, advances in telecommunications—have further enhanced these capabilities. International business elites use the news media to obtain the same information that foreign policy decision-makers use. News agencies such as Reuters and AFP know this and their economic well-being depends on their ability to provide appropriate information. Whether this information is in fact sufficient to explain international events is an empirical issue, but the incentives to provide the required information are present.

Pitfalls

1. Irregularity of news record

Despite the vast amount of information available on political events, the news media do not cover “everything.” And never will. In retrospect, the event data community under-estimated the sheer volume of "events" that occur in the world. At 1000 stories per day, Reuters initially looked like a major improvement to the ever-diminishing international coverage of the *New York Times*. But a bit of reflection will show that even 1000 stories is only a tiny fraction of all of the political events that occur in the world on any given day. *Any* news source is going to nonrandomly sample only a small amount of this activity, so the question is whether that sample is useful for a specific analytical task.

From the standpoint of mediation, two problems are particularly apparent. The first is the issue of “media fatigue” (Schrodt and Gerner 1994): the fact that the news media are much more likely to report novel events than on-going events. The initial phases of a negotiation are likely to be reported in detail, as are any dramatic changes in the negotiation such as a breakdown or an

agreement, but the day-to-day events of an extended negotiation are unlikely to be reported, even if the information is available to the press. The attention span of the media is decidedly limited.

Second, much of the day-to-day information about a mediation effort may *not* be available. While some negotiations occur almost entirely in public (though even in those cases, talks in the corridor and “walks in the woods” may facilitate breakthroughs), third-party mediation almost always involves substantial activity behind closed doors. In some instances, the true character of an agreement may not be known for decades.⁵ For example, recently declassified documents show that the resolution of the Cuban Missile Crisis involved an explicit tit-for-tat agreement concerning the removal of US missiles from Turkey in exchange for the removal of Soviet missiles from Cuba. This had long been suspected by analysts, but was denied at the time, and would not have shown up as an explicit part of any published report of the negotiations. Laboratory exercises, in contrast, can preserve the full record of the mediation.

Because of these two characteristics, event data are probably going to be more useful in studying protracted mediation situations where a number of small agreements are reached through negotiations over a long period of time. This has the advantage of preserving the novelty of the negotiation process—each new little step in the agreement will be deemed newsworthy—and, through the incremental character of the negotiation, bringing most of the detailed characteristics of the negotiations into the public record. The resolution of the Arab-Israeli dispute, which has involved a series of agreements dating back to about 1978, might work well in this regard, although it is still complicated by a series of important secret negotiations such as those leading to the Oslo agreement and assorted contacts between Israel and Arab governments prior to mutual recognition. The Balkans situation, where mediation thus far has been either very decentralized (e.g. village-by-village efforts of international peacekeeping forces) or limited in time (e.g. the Dayton negotiations), will likely be more problematic.

⁵ Some of this information can be obtained ex-post-facto from memoirs, autobiographies and other accounts, but to date machine-coding methods have not been applied to this type of material because it is usually not readily available in machine-readable form. This situation is likely to change in the near future as “electronic book” formats become more widely accepted. Nonetheless, such sources are notoriously subject to systematic biases that present the writer in the best light, and the true character of the negotiation may be distorted by various “hindsight biases” that make the eventual outcome appear inevitable. The presentation of this type of information is also less systematic than news reports, and therefore is more difficult to code with automated methods.

1.1. Regional versus Global Sources

Data services such as NEXIS and Dow Jones Interactive carry the text of literally hundreds of regional news sources. As these sources became easily available, a number of researchers thought that the regional sources could be used to fill in the gaps of the reporting of global news sources. It was assumed that services such as Reuters and AFP would contain a selective *subset* of the events reported in the regional sources, and anything really important in a regional source also would be reported by the global source.

Alas, it isn't so, at least for Reuters. Reuters and the regional sources are supplementary, rather than complementary. A number of different studies in different regions of the world have shown that Reuters reports events that are not reported in the regional sources (and vice versa). These studies include

Europe	Gerner et al 1994, Huxtable and Pevehouse 1997
Middle East	Gerner et al 1994, Schrodt and Gerner 1998.
Africa	Huxtable and Pevehouse 1997, Moore and Davis 1998
Southeast Asia	Howell and Barnes 1993

The same is true when global sources are compared: For example in his dissertation research on West Africa, Huxtable (1997) assumed that the English-language sources Reuters and BBC would focus on Anglophone states such as Nigeria, Ghana and Sierra Leone, while the French-language AFP would focus on Francophone states such as Senegal, Niger and Côte d'Ivoire. This did not prove to be the case: Reuters would sometimes pick up major events in Francophone states that were missed by AFP, and AFP would sometimes provide better coverage of Anglophone states. In some cases, it was almost possible to reconstruct the travel itineraries of individual Reuters, BBC and AFP reporters as they worked their way through West African capitals, producing a flurry of temporary detail on areas that would receive no coverage for another year.

The fact that event data is coded from a finite number of sources has been criticized by Alker (1988), among others, as privileging some interpretations of history over others. While this may be true, it is no more or less the case than the situation facing traditional studies of political behavior. Short of descending into a post-modernist quagmire where nothing can be assumed, concluded or explained, any political analysis must assume that certain events, conditions, motivations and coalitions occurred, and others did not. The traditional method of composing

accounts of political activities using a variety of documentary and autobiographical sources is one way of doing this; the processes by which text is selected for event coding is another. Each method is subject to selection bias and varying interpretation.

2. Temporal irregularity of diplomatic interactions

Real-world mediation almost invariably happens very irregularly over a very long period of time. The Israeli-Palestinian issue was originally intended by the United States to be resolved in the context of the Camp David agreements in 1978; almost a quarter-century later in 2000, the issue is still being negotiated. Agreements on resolving the issues in the Balkans don't seem to be getting off to a more efficient start, a rather sobering observation.

Delay is, of course, a standard tactic in negotiation, whether at the level of international diplomacy, civil litigation, bickering over the price of zucchini, or dealing with a five-year-old. In protracted international negotiations, delay can substantially change the underlying parameters of the situation. "Facts on the ground", to use the phrase common in the Middle East, can be created or destroyed; adversaries involved in the negotiation may change due to mortality or political circumstances, and the external international environment can be modified, perhaps as dramatically as in the end of the Cold War or the decline of the cohesiveness of OPEC. In the short term, delay may be used pressure an adversary (or the mediating party) who is faced with an upcoming election or critical parliamentary vote, or to delay past that time in hopes of getting a better deal from a successor.

While I have had little direct experience with laboratory experimental methods, my sense is that this aspect of negotiation may be one of the most difficult to accurately simulate. Experimental subjects are involved in the exercise for a relatively limited (and generally more or less known) period of time, and will probably tend to seek a resolution or escalation of the situation by the end of that time period. Maryland's widely-used ICONS simulation (see Starkey, Boyer and Wilkenfeld 1999) occurs over a more extended period of time, but even in that case, the fact that ICONS is typically done in an instructional context would mean that the incentives for pig-headed obstructionism are more limited than they are in a real-world negotiation between political antagonists. Endless negotiations going nowhere—as witnessed with Law of the Sea Conference, the Korean armistice talks, negotiations over Cyprus, and much of the Israeli-Palestinian negotiations prior to Oslo—are going to be difficult to duplicate with

most human subjects. This may be easier to duplicate in computer simulations, though providing realistic incentives for a simulated agent to delay may also be problematic.

The other place where protracted negotiation is problematic is in the analysis of the event stream itself. Given the media fatigue effect, when does a negotiation “begin” and “end”? In some instances, these time points are clear-cut—the beginnings of major power third-party negotiations are usually conspicuously heralded, as are successful agreements—but in other instances they will not be. This is particularly true for “implicit agreements”, where the parties involved reach a mutual understanding which both abide by for an extended period of time, but do not formally ratify. The SALT II agreements—not ratified, but not violated—would provide a good example of this phenomenon, but many smaller, quid-pro-quo agreements probably also exist.

3. Complexity of negotiations

The final characteristic of real-world negotiations that distinguishes them from their laboratory counterparts is the complexity of issues under consideration, and the complexity of the context of those issues.

A good illustration of this complexity is found in the territorial negotiations between Israel and the Palestinians under the Oslo agreement. The complexity starts with the fact that rather than creating a single territorial division, there are three types of control—“Area A”, with almost complete Palestinian control; “Area C” with complete Israeli control; and “Area B” which is supposedly joint control. The map defining these areas is of incredible complexity—it defines a region that Palestinians sarcastically refer to as “the Holy Land”—and in places the location of a boundary has been negotiated to within an accuracy of two or three meters, then marked by the placement of massive blocks of concrete.⁶ Negotiations over the territorial control of the former Yugoslavia appear to be no less complex.

Closely related to the problem of issue complexity is issue linkage. This will most commonly occur between the parties themselves, the “land for peace” linkage of the Israeli-Arab negotiations being a conspicuous case. The “final status” negotiations of the Oslo process

⁶ This is particularly relevant at road intersections, where control of a cross-road can impact activity such as unrestricted patrols in a much larger area. In the course of the partial Israeli withdrawal in the West Bank, heated disputes have occasionally arisen on the ground over disagreements in the location of lines on maps that differed

involve not only land and security, but a constellation of issues that include water, access to religious sites, repatriation of refugees and the release of prisoners. Linkage can also occur with the mediator, as in various security guarantees and offers of economic aid by major powers such as the United States or the European Union. U.S. monitoring and security guarantees were a critical factor in the Camp David agreements, and at the present time continually emerge in the context of negotiations between Israel and Syria.

Researcher who design experiments are, of course, completely aware that real-world negotiations are more complex than those in a laboratory or computer situation. The entire *point* of laboratory techniques is to simplify and control most of the variables in a situation in order study a much smaller number. New drugs are tested in cell cultures before they are tested in animals or humans; new chemical processes are developed in test tubes and beakers before they are deployed in factories.

In fact, both the Maryland fisheries scenario and the options in the POLI simulation are quite complicated with a variety of possible options and tradeoffs. The complexity of the Maryland protocol is probably near the limits of what student subjects could be expected to effectively deal with in the relatively short period of time available for the experiment. Consequently, when one tries to compare the laboratory and real-world events, it is quite possible that it will be more difficult to isolate a “core” negotiation from the noise of the subsidiary issues. The unconstrained (and selectively reported) event stream is going to be a lot messier than the messages found in the experiment.

A computer simulation has fewer constraints in terms of complexity, in the sense that most computer algorithms can work with 50 issues as well as they can work with 10 issues,⁷ but POLI may be near the limits of the ability of a single researcher to specify and analyze all of the parameters and contingencies required in a complex model.

Inherent indeterminacy and counterfactuals will also complicate comparison. Most negotiations reach a point where a series of equivalent tradeoffs are in play (i.e. one has reached a

by less than a millimeter and could easily be due to innocent transcription errors (to the extent that anything in this region is innocent...).

⁷ The choice of the lower bound here is deliberate: due to the “curse of dimensionality”, a method that works well with two or three issues may grind to a halt when confronted with ten, but if it can handle ten it can usually handle more. Most artificial intelligence methods have been designed from the outset to be robust against the expansion of the dimensionality of the model precisely because many traditional optimization methods are *not* robust, and consequently the methods used by Taber should scale.

plateau in whatever function is being optimized), and the choice among these is either random, or determined by contextual information that no simulation could expect to encompass, such as the location of an olive grove belonging to the cousin of a negotiator.⁸ In addition to dealing with the flat optimization surfaces generated by the parameters of an experiment, outcomes are further affected by the inherent stochastic behavior of human subjects. It is quite possible that one could end up with two outcomes that appear to be quite different despite the fact that they are almost the same in terms of the parameters underlying the mediation.

“Reality is a low probability event.” This is not an insurmountable problem, particularly if some metric is available such that the probability of any observed outcome can be computed (easier with a simulation than with an experiment), but it suggests that even under ideal circumstances, comparing the results of the different techniques may be difficult. On this issue, the event data approach is at a distinct disadvantage because it provides only the single outcome observed in the empirical world. In contrast, the laboratory methods can generate multiple outcomes based on slightly different (or even identical) initial conditions, different subjects in the experiment, or different initial values of the random number generator in the simulation.

Laboratory Protocols versus Event Data

In addition to the general set of issues noted above that are primarily related to differences between real-world and laboratory environments, there are some additional issues that relate specifically to event data.

⁸ The “kitchen table effect.” In an earlier incarnation as an environmental activist, I was involved in a protracted negotiation over the boundaries of a wilderness area in southern Indiana. After five years we had reached a political stalemate with our opponents, and made a series of compromises that removed large areas from the north-eastern part of the wilderness while adding a series of smaller areas in the south-west, the region furthest from property owned by our core opposition. We added as much of this new acreage as we dared to without attracting attention to the revised southwestern boundary. It was a hilltop here, a creek bottom there, sometimes a hollow we knew had a good campsite, but generally just guessing. Almost all of that final boundary—drawn across topographic maps on my kitchen table one bright spring afternoon—ended up in the final legislation authorizing the wilderness area. I have no doubts that the final division between Israeli and Palestinian areas in the West Bank will be constructed on much the same basis.

Potentials

1. Much of the required information can be easily coded

One of the positive aspects of this exercise has been the extent to which much of the allowable action in both the Maryland protocols and in many of the proposed features of Taber's simulation correspond very closely to the existing event data categories found in the WEIS coding system. For example, the <source> <target> structure found in POLI (Taber 1992) and EVIN (Taber 1999) is identical to the structure of event data. Most of the options used in POLI (Taber 1992:898) can be unambiguously mapped into WEIS codes; those that cannot are frequently coding categories that have been used in alternative event coding systems such as BCOW and PANDA, and as such present no problems for the event data approach (this will be discussed further in the next section).

In a similar vein, the Maryland protocol clearly works with dyadic communications between actors (including a mediator), and the offers, counter-offers, threats, promises and so forth that are message options in the experiment occur frequently in the Reuters record, and can easily be coded. (The examples in the Maryland protocol are a bit fish-specific, but adapting to these issues won't be much of a problem.) At this level, the world "seen" in the laboratory does not differ substantially from the world seen by event data.

2. The conflict-cooperation orientation of existing event data set can be modified

The differences that exist between WEIS and events found in the laboratory studies are often due to the fact that the WEIS coding scheme is very much a product of the Cold War. Most of the early work on WEIS was funded by the U.S. Dept of Defense Advanced Research Projects Agency (DARPA), and WEIS became the de facto standard for DARPA event data coding. DARPA's interest, in turn, was primarily directed at the problem of conflict early warning systems. Consequently WEIS is quite rich in its categories for *escalatory* behavior such as threats, reductions of relations and demonstrations, but very poor in categories that would describe negotiation and conflict resolution. DARPA doesn't do conflict resolution. (Somewhat counter-intuitively, the focus on early warning means that WEIS is also weak in describing the *use of force*—this is only a single category.) The other widely-used event coding system,

COPDAB, is somewhat better in that it codes separate domains of action (e.g. political, economic, military, symbolic), but COPDAB also employs a very strict conflict-cooperation continuum, as well as having far fewer categories for coding behavior than WEIS.

Prior to the advent of automated coding, this conflict-cooperation (and escalation) focus was a major constraint in applying event data to post-Cold War situations. It is still a partial constraint, because WEIS and COPDAB are the only two event coding systems that have been widely used in event data analysis. However, in the past ten years one major new system has been developed and fully implemented—PANDA—that provides extensive detail on domestic activity and also contains a number of coding categories relevant to mediation. Additional work is underway on a generalized system called “Integrated Data for Events Analysis” (IDEA: Taylor, Jenkins and Bond 1999) that is based on the WordNet synonym set (<http://www.cogsci.princeton.edu/~wn/>) and should provide a virtually comprehensive coding system for political activity.

The upshot is this: The *existing* event data coding schemes probably leave a lot to be desired for the study of mediation and the resolution of conflict, though they are probably fairly complete regarding the escalation of conflict. However, if the issue of the availability of machine-codeable news sources can be resolved (this will be discussed below), it is relatively straightforward to develop systems that will be more attuned to mediation and negotiation activity. This will involve some additional work, and one will probably need to refine those coding systems with experience—as has happened with PANDA, but not with WEIS and COPDAB once those systems went through an initial development phase (McClelland 1983)—but this task is relatively straightforward with existing technology and a bit of investment in dictionary development.

Pitfalls

1. Multiple levels of reference

One of the most conspicuous features of Maryland protocol are complex sentence structures such as the following (the italicized phrases are negotiation options allowed by the system):

While Thule appreciates Ultima's promise to *not decrease pollution control assistance*, Thule is unprepared to *avoid the use of trade sanctions*. Perhaps Thule could *avoid the use of trade sanctions* instead?

This is an example of a sentence that contains what I refer to as *multiple levels of reference* (this is my term; there is probably a formal linguistic phrase for this) within a single sentence. A parallel problem, albeit at a representational level rather than a linguistic level, can be found in Taber's simulations, which maintain a complex "working memory" and "long-term memory" that is regularly updated by the context of the events.

As I will note below, sentences of this level of complexity are not uncommon in Reuters, as shown in some of the examples of the opening sentences of Reuters stories given in the appendix. The problem is figuring out how to *code* such a sentence, and how to analyze the data from such a coding scheme. Specifically, classical event data codes only the simple subject-verb-object structure of a sentence, which I call a "primary event"

Lebanese President Emile Lahoud on Saturday urged Israel's new Labour government...

One can add compound subjects and still readily code the sentence:

Jordan's King Abdullah and Lebanese President Emile Lahoud on Saturday urged Israel's new Labour government...

But there is still more information in the indirect object about *what* is being discussed, which I call a secondary event

Jordan's King Abdullah and Lebanese President Emile Lahoud on Saturday urged Israel's new Labour government to resume suspended negotiations with Syria and Lebanon...

and also *why* it is being discussed, which I call a tertiary event:

Jordan's King Abdullah and Lebanese President Emile Lahoud on Saturday urged Israel's new Labour government to resume suspended negotiations with Syria and Lebanon to break a deadlock in regional peace-making.

COPDAB, and to a much lesser extent WEIS, coded secondary events as "issues", but I don't know of any examples where tertiary events have been analyzed at all, and only rarely are secondary events systematically analyzed. The only data available in the public domain that I know that extensively coded tertiary events was BCOW (Leng 1987),⁹ but the analysis of that

⁹ The codebook for the Global Event Data System (GEDS) indicates that it codes for some secondary and tertiary information, but as of 13 March 2000 the GEDS data were unavailable from either the ICPSR or the GEDS web site (<http://geds.umd.edu/geds/gedsdata.htm>)

data has been very limited (Leng 1993a, 1993b) compared to the amount of analysis of the primary event coding in the WEIS and COPDAB systems.

Now, to get to the level of detail found in the Maryland fisheries mediation protocol, we have to go another level or two in the complexity of the sentence. But such sentences are not unknown in Reuters. The appendix provides a sample of the first few sentences of some reports from May 1999 that involve third parties dealing with the Middle East. Most of these are not mediation per se—in fact some might be characterized as third-parties stirring up trouble—but they are fairly characteristic of the sorts of reports one finds in Reuters. (As is the one outlier report I've included in the series)

Reuters tends to use a standardized format in the introductory sentences of news stories from this region:

1. General summary of the issue
2. Direct quotation from one of the parties
3. Elaboration of the points raised in [1]

The remainder of the story gets less regular—it can involve more direct quotes, commentaries on the commentaries, background history, and so forth. But the following two lead sentences show examples of multiple levels of reference even in the first sentence of a Reuters story:

A joint communique said resumed talks between Israel and Syria and Lebanon must stick to the terms of the 1991 Madrid peace conference, which Arabs say demands that Israeli withdraw to its borders before the 1967 Middle East War in return for lasting peace.

Syria said today the U.S. veto of a U.N. Security Council motion on Israeli settlements was "the most prominent phenomenon of U.S. hostility to the Arabs and U.S. support for Israeli plans to annex the West Bank."

In the second example, we find at least four levels:

Primary:	U.S. veto of a U.N. Security Council motion
Secondary:	U.S. veto of a U.N. Security Council motion on Israeli settlements
Tertiary:	Syria said today the U.S. veto of a U.N. Security Council motion on Israeli settlements was "the most prominent phenomenon of U.S. hostility to the Arabs

4th level¹⁰ Syria said today the U.S. veto of a U.N. Security Council motion on Israeli settlements was the most prominent phenomenon of... U.S. support for Israeli plans to annex the West Bank"

By the final level, we are dealing with the Syrian interpretation of how a U.S. action on a U.N. action concerning an Israeli action reflects U.S. policy support for a related putative Israeli policy. While sentences containing 4th-level references are relatively uncommon in Reuters leads, tertiary references are quite common. 4th-level references are more common inside Reuters stories (that is, following the lead sentence) and within direct quotations. Full-story coding would probably be useful when studying many mediation processes—and relatively straightforward, even with sparse parsing (Schrodt and Gerner 1998)—but the parsing of direct quotes is generally beyond the reach of current automated-coded technology.

One can question whether statements beyond the primary or secondary level should be in an event data set, irrespective of whether they can be coded correctly. My sense, based on reading a variety of Reuters reports, is that secondary attributions that are policy statements—X commenting on Y's actions toward Z—carry some information, but not very much. When an important policy statement is being made, usually one will find a sentence in a Reuters lead reporting the primary event—"Syria denounced the U.S. veto of U.N...."—that can be coded. Furthermore, in some instances events involving attribution are a direct result of Reuters or other international media seeking comments or interviews rather than spontaneously generated events.

However, this still leads to a disjuncture between the data structure found in events—which is almost exclusively a simple source-event-target—and the more complicated structures that are used in the experimental and simulation designs. Event data analysis has been able to largely ignore secondary and tertiary information because it is generally highly aggregated (again, some of Leng's work is an exception), but this aggregation will not work if one is trying to manage detailed comparisons with the two laboratory approaches.

¹⁰ Does anyone know what word follows in the sequence 'primary', 'secondary', 'tertiary',...—my Latin isn't very good...

2. *Determining pronoun references across sentences.*

A major impediment to coding the automated coding full stories is the ability to handle pronoun references to earlier sentences. Consider the following initial sentences from a single story:

Islamic foreign ministers meeting in Burkina Faso have prepared a draft resolution condemning Israel for its blitz of Lebanon last week and demanding that it pay reparations, officials said.

On Iraq, they call for a lifting of sanctions in return for Iraqi respect of U.N. Security Council resolutions.

On Kashmir, they offer support to the region's search for autonomy.

On Kosovo, they demand an accelerated return of refugees and reconstruction of the southern Yugoslav province.

Based on the KEDS project work on the Middle East, we originally thought the issue of pronoun references would be relatively easy to solve, because pronouns in those stories usually refer to the first actor in the previous sentence, as in these examples.¹¹ However, this seems to be a quirk of some unknown (but we love'em!) Reuters editor handling the Levant, and this technique has not worked as well in other regions of the world. Determining pronoun references *across* sentences is more difficult than determining the references *within* sentences because the previous sentence may contain a number of different actors. At times the problem is simply unsolvable—the reference “Dennis Ross will meet with Ehud Barak when he goes to London” cannot be resolved on the basis of syntax alone; instead one must know *who* is going to London.¹²

3. *Sequence Comparison*

A final problem that we are likely to encounter is coming up with a suitable method for comparing the event sequences produced by the three methods. In event data analysis, the most common approach for comparing sequences is the use of weighting schemes such as Azar and Sloan (1975), Vincent (1979) or Goldstein (1992). While these appear to work fairly well in many applications, there are a couple of clear problems with scaling. First, aggregating events is

¹¹ The occurrence of the multiple “they” reference, on the other hand, is quite unusual in Middle East leads, though it may be common for the Reuters reporter in Ouagadougou, where this story originated.

¹² In this particular example, however, the pronoun reference is not needed to correctly code the sentence.

controversial: the “folk criticism”¹³ of the Azar-Sloan scale is “3 riots equals a nuclear war.” This debate goes back to the earliest event data discussions (e.g. Azar and Ben-Dak 1975; Azar, Brody and McClelland 1972) and has continued over time: see exchanges between Howell (1983) and McClelland (1983) or Vincent (1990) and Dixon (1990).

The unidimensional cooperation-to-conflict scaling might also be problematic—for example the USA-Canada or USA-Japan relationships are characterized by high levels of both cooperation and political conflict. To date this hasn't prevented the scaled data from being used successfully in a variety of studies, but that may be due in part to the fact that event data have been primarily employed to study highly conflictual situations such as the Cold War (Ashley 1980; Goldstein and Freeman 1990; Dixon 1986), antagonistic behavior in the Middle East (Azar 1972; Azar et al. 1979; Schrodt and Gerner 1997, 1998), and the conflict in Balkans in the 1990s (Goldstein and Pevehouse 1997, 1999). In all of these case, “cooperation” is largely expressed as a reduction of conflict.

This study of differing mediation styles, in contrast, may provide an instance where scaled event data does not present an accurate picture of the underlying behavior. While the *political objective* of mediation is the reduction of conflict, that will not necessarily be the short-term result. For example, one obvious problem in evaluating the effects of mediation in the Israeli-Palestinian dispute is that successful conflict-reduction by the core actors in the dispute—the Israeli government and the PLO/PNA—has frequently led to increased violence by extremists on both sides, notably right-wing settlers in Israeli-controlled areas and Islamic militants among the Palestinians.

Conclusion

This discussion has tended to focus on the problems of comparing laboratory and real-world behaviors, in part because I'm more concerned about what I can't do (e.g. code tertiary references) than what I can do (i.e. code and analyze a sequence of escalatory threats). This should not be taken as indicating that I think that the Schrodt-Taber-Wilkenfeld program¹⁴ is

¹³ As in the “folk theorem” of rational choice fame: I've heard this phrase many times over the years but I have no idea who originated it. The Azar-Sloan value for “inciting of riots” is equal to 44; “full-scale war” is 102.

¹⁴ If for no other reason, I should continue this work because it is one of the few collaborations where my name will come first in an alphabetical listing.

doomed to failure, but simply that it is going to be considerably more difficult than we initially anticipated. But in fact this may be no more difficult than bringing any other social scientific research out of the laboratory and into the field: We have simply had relatively little experience with that activity in the study of international relations.

Another possibility is that we may be just a little before our time with this project. Having thus far focused almost exclusively on what can be done with *existing* technology, I will close with a couple of suggestions about threshold technologies that might make this project easier in ten years.

Full-parsing

Natural language processing technology may have reached a point where a coding system based on full parsing (rather than the sparse parsing used in public-domain coders such as KEDS and TABARI) is appropriate.¹⁵ Full parsers are far better than sparse parsers at handling disambiguation, and can also deal effectively with issues such as subordinate phrases, negation, and the different forms of verbs. Fortunately for the efforts of the event data community, parsing is a very general problem that occurs in a wide variety of contexts, including automated indexing and retrieval, database query, speech recognition, and machine translation.

The development of parsers for English and other languages has been an active research area in computational linguistics for at least three decades; Briscoe (1996; <http://cslu.cse.ogi.edu/HLTsurvey/ch3node9.html#SECTION37>) provides a thorough survey of the issues involved.. Much of this work has been commercial and proprietary, but with the support of the European Union, the U.S. National Science Foundation and DARPA, some public domain and/or open-source parser projects are available. However, most existing automated content analysis systems do not use parsing. For example, a recent review of 15 contemporary software packages for computer-assisted text analysis by ZUMA (originators of the widely-

¹⁵ The Kansas Event Data System project has produced two sparse parsers for coding event data—the original KEDS and the newer TABARI (Text Analysis By Augmented Replacement Instructions). TABARI's parser is not significantly more complex than KEDS's, but TABARI is more flexible in how it handles individual words and their codes, and the C++ source code for the program is publicly available as open-source software running under the Linux operating system. The commercial VRA parser—"the automated-coder-formerly-known-as-FRED"—is substantially more sophisticated than KEDS or TABARI and is currently being used by a number of government projects in the US and elsewhere to produce event data for early warning systems. According to John Davies (conversation, January 2000), the GEDS project will have an automated coding capability by the summer of 2000.

used TEXTPACK system) found that only three “incorporate linguistic information” (and one of those was KEDS) (Alexa and Zuell 1999:147).

A “robust parser”—one designed to work with unedited text in general subject domains—can also deal with a variety of other tasks. For example, Carnegie-Mellon’s “LINK Parser” (<http://www.link.cs.cmu.edu/link/>)

... has a dictionary of about 60000 word forms. It has coverage of a wide variety of syntactic constructions, including many rare and idiomatic ones. The parser is robust; it is able to skip over portions of the sentence that it cannot understand, and assign some structure to the rest of the sentence. It is able to handle unknown vocabulary, and make intelligent guesses from context about the syntactic categories of unknown words. It has knowledge of capitalization, numerical expressions, and a variety of punctuation symbols.

With machine parsing, the parser would first “mark-up” the sentence, and then the coder would operate on the marked-up version, rather than from the original text. In principle, a user could therefore substitute an entirely different parser—for example, a commercial parser, a language-specific parser, or a domain-specific parser—without modifying the coding model. However, at the present there appears to be no single standard for marking syntactic structures, or even parts of speech, so the dictionaries would in fact be linked to a specific parser.

The disadvantage of the full-parsing approach is that the coding dictionaries and the coding schemes themselves would need to be substantially more complex in order to use the additional information. One would also need to deal with the issue raised earlier issue about the lack of agreement on what to do with secondary and tertiary references even if one can accurately code them.

Increased availability of text reporting political activity

The other major constraint that we face at the moment is the lack of accessibility to machine-readable reports of political activity that is comparable to our access to human-readable reports. To understand the extent to which the current situation is problematic, imagine that academic researchers had access to a machine-readable copy of Reuters (or some other comprehensive news source) with the same ease that they can access human-readable copies of *The New York Times* (the traditional source of event data).

Any decent research library has copies of *The Times* on microfilm. If we want students to examine events prior to the outbreak of the 1967 Middle East War, we just send them to the

library and they read the relevant articles at no cost. Human-coded event data can take advantage of this accessibility, but human coding is slow, inconsistent and expensive.

If Reuters or some comparable source were similarly accessible, one could simply download the appropriate texts from the library via a high-speed communications link, reformat them, and code them. Under those circumstances, there would be very little reason for standardized event data sets: Because machine coding is completely reproducible, archiving dictionaries and search commands would be equivalent to archiving the original data, just as a research project that has constructed a scale from the American National Election Survey data reports only the questions used in the scale, not the transformed data.

This hypothetical situation would also permit much greater experimentation with event data coding schemes, a development that is probably long overdue. When dictionaries of verb phrases are available—particularly dictionaries based on a comprehensive set of synonym classes such as used in IDEA, rather than the rather ad hoc category of WEIS and COPDAB—it is relatively easy to implement a new coding scheme. (This is because it is much easier to assign a code to a known verb phrase than it is to anticipate what verb phrases will occur.) Dictionaries would still need to be supplemented with the names of regional actor, but this task can be largely automated.

That is the ideal scenario: customized, regionally-specific, machine-coded event data. Unfortunately, the current situation is far from this ideal. Neither Reuters nor any other source is available cost-free from the library, but only through relatively scarce subscriptions and through a slow downloading process. Formats (for example the text indicating the story date and the text indicating the beginning of a story) are not standardized—in fact sometimes formats within a single text source are not standardized—and reformatting often proves to be a major problem of researchers undertaking a coding project. For the past two years, current reports from Reuters have not been available at all from NEXIS, the most common academic source for news reports. From the perspective of machine coding, the books are still locked up in the monastery libraries, chained to the desks, and copied by hand onto expensive parchment.

This situation is presumably temporary. Archives of most news periodicals (although not Reuters) are already available electronically through services such as NEXIS and Dow-Jones Interactive and, increasingly, on the World Wide Web. Once in electronic form, the cost of moving those archives to a high-density archival medium such as DVD is minimal, and someone will presumably find a way to profitably sell these archives to libraries, just as microfilm is

currently sold. It would be very nice if this occurred with Reuters, but Reuters has become notorious in the world of information services for resisting innovation and this is rather unlikely. More likely we will find that by the year 2010, a more up-to-date company such as Dow-Jones, CNN, AFP, Bloomberg or even Microsoft will fill this niche, and provide full-text sources going back to 1990 or so.

The availability of news reports on the web suggests another future model: automated web “spiders” that would continually search web-based sources of news reports, summarize and rewrite these using a standard vocabulary and syntax, eliminate redundant reports, and archive the results for future analysis. If this were done as a cooperative venture, much as large scale social science surveys are currently done, it would provide the missing link of a large corpus of reports on political activity unencumbered by copyright. All of the required components for such a system exist (or could be readily developed), and there is probably sufficient Web coverage of the industrialized states (but not the Third World) to provide data at least as good as what we currently get from Reuters.

Bibliography

- Alexa, Melina and Cornelia Zuell. 1999. *A Review of Software for Text Analysis*. Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- Alker, Hayward R. 1988. "Emancipatory Empiricism: Toward the Renewal of Empirical Peace Research." In Peter Wallensteen, ed. *Peace Research: Achievements and Challenges*. Boulder, CO: Westview Press.
- Ashley, Richard K. 1980. *The Political Economy of War and Peace*. London: Frances Pinter.
- Axelrod, Robert, ed. 1976. *Structure of Decision*. Princeton: Princeton University Press.
- Azar, Edward E. 1982. *Conflict and Peace Data Bank (COPDAB), 1948-1978*. Ann Arbor: Inter-University Consortium for Political and Social Research. (ICPSR Codebook no. 7767, second edition)
- Azar, Edward E., and Thomas Sloan. 1975. *Dimensions of Interaction*. Pittsburgh: University Center for International Studies, University of Pittsburgh.
- Azar, Edward E. 1972. "Conflict escalation and conflict reduction in international crisis: Suez, 1956." *Journal of Conflict Resolution* 16,2:183-202.
- Azar, Edward, R.D. McLaurin, Thomas Havener, Craig Murphy, Thomas Sloan, and Charles H. Wagner. 1977. "A System for Forecasting Strategic Crises: Findings and Speculations About Conflict in the Middle East." *International Interactions* 3:193-222.
- Azar, Edward E., Richard A. Brody, and Charles A. McClelland, eds. 1972. *International Events Interaction Analysis: Some Research Considerations*. Beverly Hills: Sage Publications.
- Dixon, William J. 1986. "Reciprocity in United States-Soviet Relations: Multiple Symmetry or Issue Linkage." *American Journal of Political Science* 30:421-45.
- Dixon, William J. 1990. "Interdependence and Cooperation in Foreign Policy Behavior: A Comment on Vincent." *International Interactions* 16:109-114.
- Esty, Daniel C., Jack A. Goldstone, Ted R. Gurr, Pamela Surko and Alan N. Unger. 1995. *State Failure Task Force Report*. McLean, VA: Science Applications International Corporation.
- Esty, Daniel C., Jack A. Goldstone, Ted R. Gurr, Barbara Harff, Marc Levy, Geoffrey D. Dabelko, Pamela Surko, and Alan N. Unger. 1998. *State Failure Task Force Report: Phase II Findings*. McLean, VA: Science Applications International Corporation.
- Gerner, Deborah J., Philip A. Schrodt, Ronald A. Francisco, and Judy L. Weddle. 1994. "The Machine Coding of Events from Regional and International Sources." *International Studies Quarterly* 38:91-119.
- Goldstein, Joshua S. 1992. "A Conflict-Cooperation Scale for WEIS Events Data." *Journal of Conflict Resolution* 36: 369-385.
- Goldstein, Joshua S., and John R. Freeman. 1990. *Three-Way Street: Strategic Reciprocity in World Politics*. Chicago: University of Chicago Press.

- 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.
- Goldstein, Joshua S., and Jon C. Pevehouse. 1999. "Serbian Compliance or Defiance in Kosovo? Statistical Analysis and Real-Time Predictions." *Journal of Conflict Resolution* 43,4: 538-546.
- Howell, Llewellyn D. 1983. "A Comparative Study of the WEIS and COPDAB Data Sets." *International Studies Quarterly* 27: 149-159.
- Howell, Llewellyn D., and Gillian Barnes. 1993. Event data for region-specific interactions: a research note on source coverage. In *International event-data developments: DDIR phase II*, edited by Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes, 45-54. Ann Arbor: University of Michigan Press.
- 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: 8-19.
- Leng, Russell J. 1987. *Behavioral Correlates of War, 1816-1975*. (ICPSR 8606). Ann Arbor: Inter-University Consortium for Political and Social Research.
- Leng, Russell J. 1993a. "Reciprocating Influence Strategies in Interstate Crisis Bargaining." *Journal of Conflict Resolution* 37:3-41.
- Leng, Russell J. 1993b. *Interstate Crisis Behavior, 1816-1980*. New York: Cambridge University Press.
- McClelland, Charles A. 1976. *World Event/Interaction Survey Codebook*. (ICPSR 5211). Ann Arbor: Inter-University Consortium for Political and Social Research.
- McClelland, Charles A. 1983. "Let the User Beware." *International Studies Quarterly* 27:169-177.
- Moore, Will H. and David R. Davis. 1998. "Ties that Bind? Domestic and International Conflict Behavior in Zaire," *Comparative Political Studies* 31: 45-71
- Reising, Uwe. 1999. "United in Opposition? A Cross-Sectional Time-Series Analysis of European Protest in Three Selected Countries, 1980-1995." *Journal of Conflict Resolution* 43,3: 317-342.
- Sankoff, David, and Joseph B. Kruskal, eds. 1983. *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison*. New York: Addison-Wesley.
- 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., Shannon G. Davis, and Judy L. Weddle. 1994. "Political Science: KEDS—A Program for the Machine Coding of Event Data." *Social Science Computer Review* 12: 561-588.
- Schrodt, Philip A., and Deborah J. Gerner. 1998. "An Event Data Set for the Arabian/Persian Gulf Region 1979-1997." Paper presented at the International Studies Association, Minneapolis.
- Starkey, Brigid, Mark A. 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. and Richard Timpono. 1994. "The Policy Arguer: the Architecture of an Expert System." *Social Science Computer Review* 12:1-37.
- Taber, Charles S. 1999. "Experiments with an Artificial Mediator: A Progress Report." Paper presented at the International Studies Association, Washington.
- 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.
- Thomas. G. Dale. 1999. *Crisis, Equilibrium and Protracted Social Conflict: The "Strange Attractiveness" of Protracted Social Conflict in Northern Ireland*. Ph.D. dissertation, University of South Carolina.
- Vincent, Jack E. 1979. *Project Theory*. Lanham, MD: University Press of America.
- Vincent, Jack E. 1983. "WEIS vs. COPDAB: Correspondence Problems." *International Studies Quarterly* 27:160-169.
- Vincent, Jack E. 1990. "Interdependence Reexamined." *International Interactions* 16:91-107.
- 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.

Appendix: Examples of Third-Party Comments in Reuters Reports on the Middle East, May 1999

(All sentences are Copyright © 1999, Reuters)

AMMAN, May 29 (Reuters) - Jordan's King Abdullah and Lebanese President Emile Lahoud on Saturday urged Israel's new Labour government to resume suspended negotiations with Syria and Lebanon to break a deadlock in regional peace-making.

A joint communique, issued after Lahoud's visit to Jordan, said resumed talks between Israel and Syria and Lebanon must stick to the terms of the 1991 Madrid peace conference, which Arabs say demands that Israeli withdraw to its borders before the 1967 Middle East War in return for lasting peace.

BEIRUT, May 29 (Reuters) - Hizbollah said on Saturday that it would continue resisting Israel's occupation of south Lebanon even if the pro-Israeli South Lebanon Army militia withdrew from the Jezzine enclave bordering Israel's security zone.

Hizbollah could enter Jezzine after an SLA pullout if it chose to, Naeem Kassem, deputy leader of the pro-Iranian guerrilla group, told reporters.

"We are a (mobile) resistance force that does not set up bases.. There is no meaning for the argument that we should not enter Jezzine because we do not represent (Lebanon's) military or political authorities," said Kassem.

JERUSALEM, May 28 (Reuters) - Israel's outgoing right-wing government expanded the borders of a Jewish settlement, angering Palestinians who on Friday called it a bid to cement Israel's hold over occupied West Bank land adjacent to Jerusalem.

Labour Party Prime Minister-elect Ehud Barak, who ousted Benjamin Netanyahu on May 17, declined comment on the expansion of Maale Adumim settlement. Labour and Palestinian officials called the action a provocation and urged the plan be dropped.

Palestinian negotiator Saeb Erekat gave John Herbst, the U.S. consul-general in Jerusalem, a letter asking for "immediate American intervention to end these destructive practices so that the attempt to put the peace process back on track can succeed".

TUNIS, May 28 (Reuters) - Libyan leader Muammar Gaddafi had talks on Friday with Sudan's First Vice President Ali Osman Mohammad Taha, a day after meeting a leading opposition figure in a bid to reconcile the Sudanese government and its opponents, Libyan state television reported.

The television said Gaddafi on Thursday had met Mohamed Osman al-Mirghani, head of the Democratic Unionist Party (DUP), one of the two main opposition groups.

Gaddafi also met Egyptian Foreign Affairs Minister Amr Moussa on Thursday in the Libyan coastal city of Sirte. Egypt has been active in reconciliation talks between the Sudanese opposition and the government.

BEIRUT, May 28 (Reuters) - Lebanese Parliament Speaker Nabih Berri said on Friday that a withdrawal from the Jezzine enclave by Israel's militia allies would not affect the struggle to drive Israeli forces from the rest of southern Lebanon.

"The withdrawals that Israel is talking about will not be covered by any (security) guarantees or arrangements that may lead to separating the Lebanese and Syrian peace tracks," Berri told a delegation of elderly people from Jezzine.

Israel offered last year to leave Lebanon if the Lebanese government gave security guarantees - suggesting a bilateral deal that would run counter to a joint Syrian-Lebanese demand for a comprehensive accord under which Israel would also return the Golan Heights to Syria.

AMMAN, May 13 (Reuters) - Hamas political leader Khaled Meshal said on Thursday the militant Palestinian group had not bowed to pressure to halt operations against Israeli targets in the run-up to elections next week.

But he said the movement, which is violently opposed to Palestinian President Yasser Arafat's interim peace accords with Israel, had been hampered by cooperation between Israel, the United States and Arafat's Palestinian Authority.

"No one can stop our right to resist the oppressive occupation or our efforts to achieve our national rights," Meshal, the most prominent Hamas figure outside the West Bank and Gaza, said in an interview in Amman.

LONDON, May 13 (Reuters) - Harrods chief Mohamed Al Fayed is so attached to his luxury London department store that he wants to stay there after he dies - as a mummy.

Egyptian-born Al Fayed has said he intends to spend the afterlife preserved in a coffin lodged in the dome on top of the store in London's prestigious Knightsbridge district, a spokesman said on Thursday.

The spokesman said Al Fayed had often spoken in a light-hearted way about being mummified, but he was "probably serious".

PARIS, May 12 (Reuters) - The French Foreign Ministry said on Wednesday it had called in Israel's ambassador to express its disapproval of a plan, which it called dangerous, to close PLO headquarters in Jerusalem.

Paris disclosed the move in releasing a letter to Faisal al-Husseini, the PLO's senior official in Jerusalem, from French Foreign Minister Hubert Vedrine.

In the letter, Vedrine said Israeli ambassador Eliahu Ben-Elissar had been told "that we disapprove of this decision which, if carried out, would violate a written commitment given in 1993 by the Israeli minister of foreign affairs".

UNITED NATIONS, May 11 (Reuters) - Secretary-General Kofi Annan said on Tuesday the Middle East peace process had languished for far too long and called for its swift resumption after Israel's forthcoming elections.

"In less than a week, Israelis will go to the polls. Whatever the outcome, an urgent task remains priority number one: courageous decision-making for peace," he said at a dinner attended by Israeli Foreign Minister Ariel Sharon marking the 50th anniversary of Israel's entry into the United Nations.

"The peace process has languished for too long. Fear has poisoned relations for far too long. Daily indignities have eaten away at families, and at society as a whole," Annan said, referring to stalled negotiations between Israel and the Palestinians.

He called for "peace with justice for all, so that the aspirations of one side are not achieved at the expense of the rights of the other."

CAIRO, May 11 (Reuters) - The Arab League on Tuesday condemned Israeli Prime Minister Benjamin Netanyahu's decision to close offices in Orient House, the PLO's headquarters in East Jerusalem.

"This step can be considered another aggression on the legitimate rights of the Palestinians, and a violation of the previous commitments to which Israel agreed in the peace process," the Cairo-based league said in a statement.

It said the step violated international laws on Jerusalem but it had full confidence the Palestinians would remain "steadfast" in achieving their rights, including "setting up an independent state with Jerusalem as its capital."