Empirical Indicators of Crisis Phase in the Middle East, 1979-1995

Forthcoming in the *Journal of Conflict Resolution*: (October 1997)

Philip A. Schrodt and Deborah J. Gerner

University of Kansas

Authors' Note: An earlier version of this paper was presented at the 1995 annual meeting of the Peace Science Society, Columbus, Ohio (October). We received valuable suggestions on revisions from Phillip A. Huxtable and Jon C. Pevehouse as well as from the reviewers. This research was funded in part by the National Science Foundation through grant SBR-9410023 and the University of Kansas General Research Fund grant 3500-X0-0038. Address correspondence to Philip A. Schrodt, Department of Political Science, University of Kansas, 504 Blake Hall, Lawrence, KS 66045, USA; email: p-schrodt@ukans.edu.

ABSTRACT

A number of studies of crisis behavior—for example, research using the Butterworth, SHERFACS, and CASCON data sets—assume that political behavior goes through a series of clear "phases" characterized by distinct patterns of interactions. Correct identification of these phases is important in crisis forecasting and in the application of mediation techniques such as preventive diplomacy. To date, empirical work with these data sets has identified crisis phases contextually (by human coders) rather than through any systematic procedures.

This article uses several statistical techniques to identify and analyze phases in an event data set measuring the political behavior between eight Middle Eastern actors—Egypt, Israel, Jordan, Lebanon, the Palestinians, Syria, the United States, and USSR/Russia—for the period July 1979 to June 1995. Consistent with narrative accounts, factor analysis identifies the Israeli-Palestinian conflict and the war in Lebanon as the two most important features of the data; other factors reflect inter-Arab and major-power relationships. Discriminant analysis can distinguish a set of human-coded phases with about 90% accuracy; stepwise discriminant achieves 70% accuracy using data from only 12 of the 54 directed dyads. Finally, K-Means cluster analysis identifies five distinct phases that align fairly well with the human-coded phases, particularly in the first half of the time period. The paper concludes with observations on how these analytic approaches might be applied to the problem of crisis early warning.

Introduction

In recent years, the topic of early warning—moribund for about a decade after substantial research funded by the U.S. Defense Advanced Research Projects Agency (DARPA) in the late-1970s (e.g., Singer and Wallace 1979; Choucri and Robinson 1979; Hopple, Andriole, and Freedy 1984)—has received renewed attention in the international relations literature (Rupesinghe and Kuroda 1992; Gurr and Harff 1994; Gurr and Harff 1996). With the end of the perceived threat of Communist exploitation of ethnic divisions, the liberal-democratic military powers—the United States, Britain, and France—are less inclined to intervene unilaterally in local or regional disputes. The international community has instead increasingly relied on multilateral responses, including the recycling of Cold War organizations (NATO in the former Yugoslavia and the United Nations generally), *ad hoc* initiatives (Iraq-Kuwait, Rwanda, Bosnia), and the use of existing non-military organizations in a peace-keeping role (ECOWAS in Liberia).

This dependence on multilateral responses enhances the attractiveness of early warning in two ways. First, there is general agreement (Cahill 1996; Crocker and Hampson 1996; Lund 1996; Schmeidl 1997) that a conflict in its early stages can often be contained by either limited force or diplomacy backed with the threat of force or other international sanctions. Second, multilateral actions require substantially longer to orchestrate than did the rapid responses of a superpower or a Cold War alliance. This has led to significant interest by international organizations in early warning (e.g., Boutros-Ghali 1992; Dedring 1994; Alker, Gurr, and Rupesinghe 1995; Mizuno 1995).

In addition, improvements in communication and computer technologies have changed dramatically the quantity and timeliness of the information available for use in early warning. Material relevant to political early warning is available from the commercial efforts of Reuters, *Agence France Press*, 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 (King 1996;

http://www.reliefweb.int). The Internet and data providers such as NEXIS provide a quantity of real-time information far exceeding that available to the Central Intelligence Agency (CIA) and KGB during most of the Cold War period.[1] Similarly, inexpensive desk-top computers now surpass in speed and memory most of the computers available to national intelligence agencies until the mid-1980s. These machines can process text-based electronic communications of news organizations, IGOs, and NGOs directly without the delays caused by labor-intensive human event coding. Whether this massive quantity of information can be effectively *analyzed* is another issue—this is the crux of the early warning challenge—but a researcher working with public domain sources in the late 1990s has access to dramatically more real-time information and data processing capability than she or he would have had available even a decade earlier.

In conjunction with this increased interest in early warning, researchers are also using a more sophisticated model of "crisis" than that employed earlier. Most of the quantitative DARPA studies worked with continuous indicators of crisis. For example, the DARPA-funded Early Warning and Monitoring System (EWAMS) system

> provides a comprehensive profile for the particular [dyadic] relationship in terms of conflict and other probabilities (for monthly data) and indicator readings (for total activity and other standard system indicators in their raw and transformed or Z-scores versions, for daily, weekly, monthly, quarterly or yearly time intervals selected by the user). (Hopple 1984, 52)

In contrast, a number of contemporary studies of crises assume that political behaviors go through a series of phases that are *qualitatively* delineated by an emphasis on different sets of behavior. In the statistical literature, crisis "phase" has been coded explicitly in the Butterworth international dispute resolution data set (Butterworth 1976), CASCON (Bloomfield and Moulton 1989, 1997) and SHERFACS (Sherman and Neack 1993).[2] Describing the early CASCON work, Sherman and Neack explain that

> conflict is seen "as a sequence of phases." Movement from phase to phase in a conflict occurs as "the factors interact in such a way as to push the conflict ultimately across a series of

> *thresholds* toward or away from violence" (Bloomfield and Leiss 1969). Characteristics of disputes can be visualized as the timing and sequencing of movement between and among phases. Processes of escalation of violence, resolution or amelioration of the seriousness (threat of violence-hostilities) and settlement are identifiable through the use of phrase structures. (Sherman and Neack 1993:90)

CASCON and SHERFACS, for instance, code six phases: "dispute phase," "conflict phase," "hostilities phase," "post-hostilities conflict phase," post-hostilities dispute phase," and "settlement phase."

In the policy literature, crisis phase has emerged as a key aspect of the "preventive diplomacy" concept, because of the assumption that diplomacy can be more effective in the early stages of a crisis (e.g., before the outbreak of military hostility) than in later periods (Rupesinghe and Kuroda 1992; Lund 1996; Bloomfield and Moulton 1997). To illustrate this argument, Lund (1996, 38-39) outlines a series of crisis phases ranging from "durable peace" to "war" and emphasizes the importance of preventive diplomacy during the "unstable peace" phase. In situations where preventive diplomacy is not an option, crisis phase may still be of utility in providing early warning of, for instance, large-scale refugee movements. Depending on the crisis phase, a localized outbreak of military action may be contained without generating large numbers of refugees, or it might rapidly spread, requiring the need for an international response. Finally, much of the literature on ethnic conflict assumes that militarized ethnic disputes such as those found in the former Yugoslavia, the former Soviet Union, Rwanda, Sri Lanka, and elsewhere evolve through a series of relatively predictable phases (Alker, Gurr, and Rupesinghe 1995; Leatherman and Väyrynen 1995).

The crisis phases identified in the Butterworth, CASCON, and SHERFACS data sets have all been assigned retrospectively by human coders. While this type of coding is obviously necessary in the early stages of a new concept's development, it presents two problems. First, when the coding of a crisis phase is dependent on human judgment, the *de facto* definition of the phase is likely to drift over time. This can happen as a single coder becomes more familiar with the data and is also likely during attempts to transfer the definition of a crisis phase across projects. Consequently, the crisis phases coded in two different data sets may appear to have conflicting implications because the coders were, in fact, working with disparate definitions. In contrast, the statistical identification of phases—combined with the machine-coding of event data (Gerner et al. 1994)—should make it possible to code crisis phase consistently and efficiently within a variety of contexts and from an assortment of different news sources.

Second, the tendency of human analysts to impose order on political events means that in some instances human coders may identify phases that do not correspond to information reported in the data set. If the human coder correctly identifies the phase, but unconsciously makes that assessment based on exogenous knowledge rather than the variables in the data set, any model that attempts to predict the phase using those data will be subject to specification error. Conversely, if the human coder has incorrectly identified the phase, any statistical estimates made with the data will be biased. We suspect that human-coded phase identification contains both types of error.

This paper identifies and analyzes the phase structure of political events involving Egypt, Israel, Jordan, Lebanon, the Palestinians, Syria, the United States, and the Soviet Union/ Russia for the period 1979-1995 using event data.[3] This region—and the event data set describing it—contains several inter-linked disputes. The two dominant political themes have been the Israeli-Palestinian conflict and the Lebanese civil war, both of which have gone through phases of hostility and mediation. In addition, there were other key focal points, such as US efforts to resolve the larger Arab-Israeli dispute and spin-off interactions resulting from both the Iran-Iraq War and the Iraq's invasion of Kuwait. This extensive foreign policy activity presents a realistic challenge to any effort to identify crisis phase through statistical indicators because of the quantity and variety of material.

Methodology and Data

One of the most substantial difficulties in using event data to analyze behavior in an international subsystem is the sheer quantity of information available. An eight-actor system such as the one we are analyzing has 56 directed dyads (the actions of X toward Y, indicated as X>Y), each with the potential to contribute to the overall interaction patterns of the system. However, much of this activity is inter-correlated because it is generated by a small set of political issues. In order to characterize the behavior of the system, therefore, we need to ascertain the underlying issues that are generating the observed actions, then identify the contributions of various dyads to determining those behaviors. If the concept of crisis phase is correct, we would expect that the observed behaviors would fall into distinct patterns over time and that we could determine the event behaviors that are characteristic of any phase. In addition, if the behaviors determining a phase are distinct and are reflected in event data, it should be possible to identify those phases inductively by looking at the data themselves, without prior knowledge of the phases.

The approach we use to analyze the system is to examine the behavior of the vector

$$[AB, AC, AD, \ldots, AH, BA, BC, \ldots, BH, CA, \ldots, HF, HG]_t$$

where A, B, ..., H are the actors in the system and XY_t is the Goldstein score of events directed from X to Y, aggregated over a month.[4] We define the "behavior" of the system as the path this vector traces over time in an N-dimensional space. A crisis "phase" is indicated by a region in the vector space where points cluster over time. Empirically, a phase typology characterizes the behavior of the system provided that most of the time these behavioral points are found inside a few distinct clusters, with brief transitions between these clusters.

Figure 1 is an idealized illustration of this process for the World War II period, using just two hypothetical dimensions: "talking versus fighting" and "local versus global involvement." The years prior to 1936 involved little violent conflict. The system then shifted to a series of militarized crises during the period 1936-38 and erupted into a full-scale European war in 1939-40. After a lull in the early part of 1941, the conflict spread first to the USSR, then to the Pacific;

1942-1944 was characterized by a global war. In 1945, this war ended, first in Europe and then in the Pacific, but post-war politics remained global rather than returning to the unilateralism/isolationism of the pre-war period. The 1946-47 cluster continued to characterize the system for most of the Cold War period, with occasional departures from this cluster to take in the Korean War, the Suez War, the Cuban missile crisis, and other militarized events.

Figure 1 about here

Obviously an analysis using event data will be more complicated than this due to the problem of aggregating dyadic behaviors, the existence of multiple issues determining those behaviors, and the fact that actual political behavior is considerably noisier than the short-answer-exam summary of international politics in the 1930s and 1940s presented above. Nevertheless, if interactions in a system can be characterized by a phase typology, and if the behaviors defining those phases are captured by event data, it should be possible to determine behavioral phases using statistical clustering techniques.

In this paper, we first use factor analysis to determine whether it is possible to reduce the dimensionality of the vector describing the system by using the correlations between interactions of the various dyads. We then examine the clustering problem both deductively and inductively. In the deductive analysis, we begin by positing a set of behavioral phases in the Middle East political landscape during the 1979-1995 period, based on our contextual knowledge of the region. Using discriminant analysis, we ascertain the extent to which the dyadic behavior reflected by event data can predict these human-identified phases. Examination of the discriminant space also provides insights into the types of behaviors that are most important in determining the phase.

In the inductive part of the research, we attempt to discern system phases directly from the observed data by using cluster analysis (Aldenderfer and Blashfield 1984, Bailey 1994), rather than setting the phases *a priori*. In other words, we look for clustering in the data themselves rather than externally imposing any order upon them. Once we determine these, we look at whether the system tends to remain in a cluster for a period of time, as postulated by the phase model, then try to interpret the political meaning of the clusters based on our contextual knowledge of the situation. We also briefly consider whether there are early warning indicators that indicate when the system is ready to shift from one cluster to another.

The data used in this study include all interactions among Egypt, Israel, Jordan, Lebanon, the Palestinians, Syria, the United States and the Soviet Union/Russia (except for USA>USR and USR>USA), as reported by Reuters News Service lead sentences downloaded from the NEXIS data service.[5] This gives us a total of 54 directed dyads with 192 monthly totals in each dyad.[6] We machine-coded these data according to the WEIS coding scheme (McClelland 1979) using the Kansas Event Data System (KEDS), a computer program that generates event data from machine-readable reports. The KEDS coding dictionaries and the data used in the analysis can be downloaded from the World Wide Web site [http://www.ukans.edu/~keds].[7]

KEDS operates by first doing some simple linguistic parsing of the news reports (for example, it identifies the political actors in the text, recognizes compound nouns and compound verb phrases, and determines the references of pronounce). It then employs a large set of verb patterns to determine the appropriate event code. KEDS agrees with human coding for around 85% of its event assignments when coding WEIS events for the Middle East and is completely replicable. The KEDS program is described in Gerner et al. (1994) and Schrodt, Davis, and Weddle (1994); Schrodt and Gerner (1994) and Huxtable and Pevehouse (1996) discuss the validity of this approach to generating event data.

Prior to doing the analysis, we assigned the months in the data set to the phases identified in Table 1. These phases reflect the dominant issues and activities affecting the region and are not intended to parallel exactly the crisis phases of the CASCON and SHERFACS data sets. They do, however, mirror many of the characteristics of the CASCON and SHERFACS phase structures, notably the movement between periods of violent conflict and periods of dispute resolution.

Table 1 about here

Factor Analysis

Our first objective was to see whether we could decrease substantially the dimensions of the N-actor system's behavior without significant loss of information. Reducing dimensionality is possible when there are consistent correlations between some of the dyadic behaviors in the system. There were at least three substantive reasons we anticipated this might occur.

First, there is considerable policy coordination between some of the states. In the extreme case, since the late 1980s Lebanon's foreign policy has been greatly influenced by Syria. Lebanon's reported behaviors generally mirror those of Syria, so knowledge of Syria's position toward another actor may by itself provide sufficient information to predict Lebanon's policy toward that actor. A less extreme example is conventional policy coordination: For instance, during the Reagan years US and Israeli policies closely paralleled each other on most issues affecting the region, as did those of Syria and the USSR.

Second, all of the states in the system are reacting to at least some of the same events: Israel's invasion of Lebanon, Syria's eventual establishment of military hegemony in Lebanon, the Palestinian *intifada*, and the Madrid and Oslo negotiations, among others. To the extent that states share similar policy positions, they may react to these external stimuli in comparable ways.

Finally, it is likely that some of the actors in our set of states have very little influence on overall activity within the system. Two candidates are Egypt—which was diplomatically isolated during much of the period we are analyzing due to its signing of the Camp David agreements with Israel—and Jordan, which is comparatively small and neither initiates nor receives many events. It is possible that a simplified model of the system's behavior can ignore these actors without a significant decrease in predictive value.

While there are a variety of methods that can be used to reduce dimensionality through correlation, we focused on the oldest and most well-understood: factor analysis (Kim and Mueller 1978). Factor analysis creates clusters of variables (in this case, the vectors of monthly net cooperation scores for each directed dyad) based on their mutual correlations. Consequently, it allows us to evaluate both the extent of the correlation between dyadic behaviors and the extent to which the dimensionality of the system can be reduced. If foreign policies are primarily determined by exogenous events—whether the foreign policy of another actor or the interactions of other dyads (e.g., the Israeli-Palestinian conflict)—then those policies should show up as a distinct factor. Furthermore, the political content of a factor cluster should be apparent from the dyads with which it is most strongly correlated. Second, if a state is not actively involved in the system (or is uninvolved with some subset of issues), this will be evident from an absence of correlation between its behavior and the factors that dominate the system.

We factor analyzed the monthly aggregated scores for the 54 directed dyads using principal components, then rotated the original factors using the varimax criterion to minimize the number of variables that have high loadings on any given factor.[8] The results of the factor analysis are presented in Table 2, which shows the factors that have the five highest eigenvalues and the directed dyads with which those factors are most strongly correlated. As anticipated, the two factors that explain the highest amount of variance are those associated with (a) Israeli and Syrian involvement in Lebanon and (b) Israeli-Palestinian dispute.[9] The third factor appears to emphasize the dyads involved in the Camp David peace process, the fourth reflects the US involvement in the Lebanon dispute, and the fifth involves Jordan's interactions with Syria and Lebanon. This last factor is contrary to our expectation that Jordan is relatively unimportant in regional behaviors.

Table 2 about here

In the varimax rotation, most of the dyads correlate strongly with only a single factor; the exceptions are ISR>USA and ISR>PAL (correlating with factors 2 and 3) and LEB>USR (correlating with factors 3 and 4). Thirteen dyads have no correlations of 0.20 or higher with the first five factors. All of these dyads involve either the USSR/Russia, which during the period studied was a relatively minor actor in the region, or—as we had anticipated—Egypt.

The first five factors explain only 27.8% of the total variance in the original data; by the usual standards of factor analysis, this is quite small. There are 21 factors with eigenvalues greater than 1.0 (the conventional rule-of-thumb for significant factors) and the factors beyond the five shown in Table 2 exhibit a long and very gradual decline in the variance explained. Most of those smaller factors appear to be picking up the idiosyncratic behavior of one or two dyads.

Based on this analysis—and contrary to our expectations—we conclude it is not possible to reduce the behavior of this system to a small number of factors. The high intrinsic dimensionality we find in the data may be due in part to the wide variety of behaviors shown in this long time series and the political diversity of the dyads. In some earlier experiments using only the dyads involving Israel and the Palestinians plus LEB>SYR, LEB>PAL, SYR>ISR and SYR>LEB as sources of behavior between 1982 and 1993, we found that four factors explained about 70% of the variance (Schrodt and Gerner 1995). Factor analysis may therefore be useful as a data reduction technique in systems less diverse than the one we've considered here.

Discriminant Analysis

Discriminant analysis is used to classify cases in a data set into a group of previously known, nominal categories based on linear combinations of the values of the interval-level independent variables describing each case (Klecka 1980). In the discriminant analysis in this study we use the behaviors of each dyad to classify the monthly data into the phases we assigned in Table 1. In other words, the "case" is a month of political activity; the "category" is the *a priori* assignment in Table 1 of each month to a phase, and the variables describing each month are the aggregated Goldstein scores reported for the 54 directed dyads. A month that shows activity typical a phase's general behavior will be located close to the center of the cluster containing the months of that phase—near the "centroid" of the phase in the discriminant space—while a month with atypical behavior (such as an outbreak of violence in an otherwise peaceful phase) will be located far from the centroid and will probably be misclassified into another phase.

Discriminant analysis is one method of determining whether the information required to describe the various phases of political activity is actually present in the event data set. If the analysis using the event data available for the 54 dyads cannot assign most months to the correct phase, then we must conclude that those human-identified phases were actually determined by information beyond what is available in the monthly event data, such as the hindsight of the analyst or trends aggregated over more than a month.[10]

On the other hand, if discriminant analysis can classify most cases successfully, the characteristics of the discriminant space and the locations of the clusters within that space provide additional information about the data set as a whole. For example, the location of the centroids in the discriminant space usually provides some insight about the attributes that distinguish each phase from the others.

We first ran the discriminant analysis with all of the dyads; we then did a stepwise analysis to determine whether the phases could be identified by looking at only a small number of dyads. The results of these experiments are presented in Tables 3a and 3b; Figures 2a and 2b show the first three dimensions of the discriminant space.

Tables 3a and 3b about here

As Table 3a indicates, when we include all of the dyads, the discriminant analysis differentiates the phases with a high degree of accuracy: 90% of the months are classified correctly and most of the errors are plausible (e.g., *Oslo* for *Madrid*; *Madrid* for *Taba*). This means we can successfully determine phase assignments from the patterns of monthly dyadic behaviors without additional information. Unlike the factor analysis results, the discriminant analysis concentrates most of the explanatory power in the first three dimensions of the discriminant space: these explain about 74% of the variance in the classification.

The stepwise discriminant analysis reported in Table 3b chose the following 12 dyads (in order of selection; the number in parentheses is Wilks' lambda[11]):

ISR>PAL (.636)	LEB>ISR (.427)	PAL>UAR (.324)
ISR>JOR (.250)	UAR>USA (.203)	PAL>LEB (.173)
JOR>ISR (.146)	UAR>ISR (.127)	SYR>LEB (.110)
UAR>JOR (.094)	ISR>LEB(.081)	JOR>USA (.071)

With 70% of the cases classified correctly, the stepwise discriminant is considerably less accurate than the system using all of the dyads, although 45.4% of the incorrect classifications occur into phases immediately before or after the correct phase. The stepwise selection appears to pick a single dyad to represent each type of dominant behavior in the system—for example, ISR>PAL is included but not PAL>ISR—and eliminates dyads that are correlated with those selections through reciprocity or coordinated policies. Unsurprisingly, nine of the dyads included in the stepwise discriminant analysis involve either Israel, Lebanon, or the Palestinians.

Figures 2a and 2b about here

Figures 2a and 2b show the first three dimensions of the discriminant space; in both cases, the diagrams are based on the analysis of all 54 dyads rather than the stepwise analysis. In Figure 2a, the first (horizontal) dimension discriminates the phases *chronologically*: the centroids are in temporal order from left to right, except for the swapping of *Camp David-Lebanon* and *intifada-Kuwait*. This is particularly interesting given that none of the variables contain explicit chronological information; the discriminant analysis is instead picking up the changes in behavior over time. The interpretation of the second (vertical) dimension in Figure 2a is less clear; it may serve primarily to differentiate the *Camp David* and *Lebanon* phases, which strongly overlap in the first dimension.

Because the *intifada* is one of the most conspicuous features of the data set, a puzzling aspect of Figure 2a is the location of the *intifada* cluster, which is thoroughly intermixed with the *Madrid* and *Oslo* clusters. As Figure 2b shows, this is an effect of the projection: the third dimension clearly separates the *intifada*, *Madrid*, and *Oslo* phases and generally seems to reflect a violent conflict versus conflict resolution dimension.

Cluster Analysis

The discriminant analysis demonstrates that we can differentiate previously identified political phases with a high degree of accuracy using information about the dyadic behaviors in the system. This means that the phases are statistically distinct and their identification does not require additional information, such as hindsight bias, that is available only to the human coder. However, the discriminant results do not necessarily mean that the phases we have identified *a priori* are the same as those that would arise naturally from clusters of data points in the 54-dimensional space.

In order to determine whether clusters are actually present in the data, we used the SPSS K-Means clustering algorithm (Norusis 1994) with the Euclidean metric

$$\sum_{i=1}^{54}(x_i - y_i)^2$$

as the distance between points. The K-Means algorithm starts by finding K cases that are widely separated in the vector space; these are used as the initial seeds for generating the clusters. The algorithm then assigns each of the remaining cases in the data set to the cluster whose center is closest to the case. When this is completed, the center of each of the K clusters is recomputed and all of the cases are again assigned according to their proximity to the cluster centers. Because those centers are now computed from the assigned clusters rather than from the location of the seed case initially used to establish each cluster, this will cause some changes in the cluster assignments. This process is repeated until the cluster membership no longer changes; the algorithm usually converges after a small number of iterations.

Because K-Means starts with cluster centers that are widely separated, some of the initial centers will be outliers and their clusters will contain only one or two points after the iterations have been completed. Consequently, we used a two-stage process: In the initial iteration we used a relatively high number of centers (16), then identified the centers of the large clusters. These large clusters were quite distinct: the large clusters contained 20 to 50 points, whereas the remaining clusters contained fewer than five points. We then used only the centers of the large clusters as the starting points for a new clustering that assigned the outlying points to the large clusters.[12]

Figure 3 about here

The results of the clustering using the 54 directed dyads is shown Figure 3. The five clusters identified by the K-Means procedure are labeled "A" through "E"; the presence of a bar indicates that a monthly data point was assigned to this cluster. The vertical lines show the phase divisions that we defined in Table 1. The graph has been smoothed using a four-month moving mode; the "X" cluster contains those points where the assignment is changing so quickly that the mode is undefined.[13] Note that because of the modal smoothing, an abrupt change of the form XXXYYYY (e.g., the *Camp David-Lebanon* transition and the *Taba-intifada* transition) shows up in the smoothed data one month before the actual transition.

Three of the clusters in the 54-dyad analysis correspond closely to the phases we had identified in Table 1. Cluster A corresponds to the *Camp David* period. It is quite uniform, although it shifts into a new phase several months before Israel's 1982 invasion of Lebanon, probably reflecting the political instability that preceded Israel's attack. This is followed by a short phase—Cluster B—that corresponds to Israel's initial invasion of south Lebanon and the siege of Beirut. The B cluster is found at only one other point, near the end of Israel's withdrawal to south of the Litani River in 1985. The system goes back into the Cluster A/*Camp David* pattern for most of the period when the US-led multinational force was in Beirut, then shifts into a new Cluster C pattern that is maintained until the end of the *Lebanon* period. The Taba period is dominated by a single cluster assignment, D, with a couple of jumps into clusters

A and C. The *Taba* phase ends abruptly at the expected transition point corresponding to the outbreak of the *intifada* and a new cluster assignment, E, develops.

At this point, the phases in Table 1 and those identified by the clustering algorithm part company. The *intifada* phase determined by Cluster E is much shorter than we anticipated; after this, the system shifts back to Cluster D, seen earlier during the *Taba* phase. The Cluster D assignment is maintained without any break through the transition between the *intifada* and *Kuwait* phases. In the unsmoothed data, there are short-term changes around the remaining phase transition points we'd identified, but these periods do not form distinct clusters. The system returns to Cluster E near the end of the *Madrid* period, which is consistent with an upsurge of violent incidents between Israelis and Palestinians during this time. The *Oslo* period is characterized by an unusually frequent pattern of cluster transitions, but does not form a distinct cluster.

We were puzzled by the inability of the K-Means algorithm to identify the Madrid and Oslo peace processes as distinct periods. We speculated this might have occurred because we initially defined our phases primarily with respect to the Israeli-Palestinian conflict, whereas the data set includes behavior from a broader set of countries. We therefore did an additional analysis looking only at the dyads involving Israel or the Palestinians as a source or target.

Figure 4 about here

These results are shown in Figure 4. The analysis found seven clusters, although neither of the two new clusters correspond to the *Madrid* and *Oslo* phases. Clusters A*, B* and C* delineate the *Camp David* period, the Lebanon invasion, and Israeli pre-1985 occupation of Lebanon in a manner similar to that of the earlier analysis. A single cluster (D*) extends from Spring 1989 until Fall 1992, again covering the *intifada-Kuwait* transition without a break, and Cluster F* distinctly delineates the intense initial months of the *intifada*, much as Cluster E had done in the first analysis.

There are two new clusters in Figure 4. Cluster E* is confined entirely to the *Taba* period—in the modal smoothing it is reduced to a single point but in the unsmoothed graph six dispersed months during the Taba period are assigned to this cluster. Cluster G* appears to correspond to intermediate levels of Israeli-Palestinian violence before and after the *intifada* cluster. The *Madrid* and *Oslo* periods still do not receive separate cluster assignments, although the unsmoothed behavior in these two phases jumps around between all of the other clusters rather than settling into a single pattern. In the unsmoothed data this instability is also found during the *Taba* phase; in fact, the cluster we have labeled *Taba* is probably a post-*intifada* phase that also characterizes the *Taba* period.[14]

Most of the clustering results appear to be consistent with our assumption that a system will begin to show instability in the behavior space prior to experiencing a phase shift. In the analysis of the 54 dyads, only the *intifada-Kuwait* transition—which was completely exogenous to the Levant—does not show some fluctuation in phase assignments prior to the transitions. This instability is only a necessary, not a sufficient, condition, but as a precursor to the phase transitions it offers the possibility of some form of early warning. The *Madrid-Oslo* period in the Israeli-Palestinian clustering is characterized by almost continuous movement between clusters and it may be possible to use this information to characterize that phase. To date we have looked only at the assignment of points to the nearest cluster center. More sensitive numerical measures—the obvious refinement is measuring a point's distance to various cluster centers—may provide some useful early-warning measures.

Analysis and Conclusion

The results of these experiments are generally encouraging for the prospects of studying phase structures systematically using statistical methods. In this section, we will briefly assess where we expect to take each of these techniques as we move toward early-warning and phase-assessment models.

Factor analysis appears to be the least useful approach, at least for this region of the world. The clusters of variables identified by the varimax rotation are politically plausible and correctly reflect what we view to be the two dominant political features of the data set (as well as identifying a politically plausible set of less-important dyads). However, the first five factors explain only about one-quarter of the variance in the data and the slope of the scree plot after those five factors is very shallow, so a large number of additional factors would be required to explain additional variance. Consequently, factor analysis does not appear very promising as a data reduction method; one should instead use the original behavioral variables.

The results of the discriminant analysis, on the other hand, were reassuring—the event data coded from Reuters lead sentences were sufficient to distinguish our *a priori* phase assignments with a high degree of accuracy and, furthermore, the discriminant space involved politically plausible dimensions. The information required to identify these phases is, in fact, present in the event data. Unlike the factor analysis, the discriminant analysis quite substantially reduced the dimensionality of the behavior, focusing about three-quarters of the discriminating power into three dimensions of roughly equal importance. *If used judiciously*, stepwise discriminant might be helpful as a technique for reducing the number of dyads that need to be monitored in the system in order to assess its phase and provide warning of phase transitions. Because any stepwise procedure is highly affected—aye, thrives on—collinearity, we do not suggest simply accepting the stepwise results as given. However, stepwise discriminant analysis might provide a means of simplifying the data requirements of an event monitoring system if it was combined with additional knowledge such as information on the consistency and reliability of the newswire reports used to generate the event data for various dyads.

Finally, the K-Means clustering worked surprising well in two respects. First, in a number of instances, it either indicated the same phases we assigned initially or else identified plausible alternative phase transitions (e.g., an earlier end to the *intifada* phase than we predicted and the "sub-*intifada*" phase in the Israel-Palestinian analysis). Given that the cluster analysis was completely inductive, used only the aggregated dyadic behaviors, and had no indication of the overall structure of the political behavior, it worked quite well. The consistent pattern of cluster instability prior to most of the phase transitions also appears promising as a possible early-warning indicator.

Additional Approaches

The analysis presented here works in a continuous variable space by aggregating events using the Goldstein scale. Earlier work (Schrodt and Gerner 1994) shows that the Goldstein-scaled values generated from events coded from Reuters reflect events on the ground fairly well; at the same time it is unlikely that Goldstein scaling—which was developed simply to give a rough estimate of the magnitude of political behavior on a conflict-cooperation continuum—is optimal for deriving factors, clustering, or predicting phase transitions.

An obvious alternative to the Goldstein scale is to use the counts of the nominally-coded events themselves. The reason we have not done this is that it expands the dimensionality of the system's behavior by a factor of 22 if one aggregates to the level of 2-digit WEIS scores and by a factor of 66 if one goes to 3-digit WEIS scores. This would require analyses where the number of variables was considerably larger than the number of data points, not to mention the computational nightmare of dealing with 1188- or 3564-dimensional matrices, most of whose entries are zero.

The Goldstein scale captures the most obvious feature of the event data—the conflict-cooperation dimension. In Table 4, we show the results of a factor analysis of the event counts in the 2-digit WEIS categories for two of the densest dyads: ISR>LEB and ISR>PAL.[15] There is a clear distinction between the amount of variance explained by the first factor and that explained by the remaining factors, and the correlation between the first factor and the Goldstein value is quite high. The results in Table 4 indicate two characteristics of the relationship between the aggregated Goldstein scores and the disaggregated frequencies of the WEIS event categories. First, over half of the variance in those counts is already reflected in the Goldstein score. Second, there is only a *single* conspicuous dimension to the event counts—one does not find, for instance, distinct and equally important dimensions for conflict and cooperation. These results suggest that an analysis of the event counts would probably not lead to dramatically different results than those found in the analysis of the Goldstein scores, at least for this international subsystem.

Table 4 about here

While the discriminant analysis found that chronological time was the single best discriminating factor, we have yet to include time explicitly as a variable in our analyses. Time is probably more important as a surrogate for exogenous events in the system, such as the end of the Cold War, than as an actual variable; however, it might also reflect the evolution of some interaction patterns. Adding time to the cluster analysis might eliminate some of the instability seen in the existing phase assignments and could give a more systematic method of smoothing than is provided by the moving mode we used here.

The Euclidean metric used in this analysis is only one of many measures that could be used for clustering. An obvious alternative approach would be to assign different weights to the various dyads. This weighting could be based on the total number of reported interactions or, conversely, it could adjust the interactions based on the overall level of reports in Reuters (on the assumption that a single reported interaction in a poorly-covered dyad such as Syria-Jordan is worth more than a number of interactions in a well-covered dyad such as Israel-USA).

Implications for Early Warning

To the extent that the Middle East is typical, this analysis has at least three implications for the use of event data in early warning research. First, it is clearly possible to analyze the behavior of an N-actor system taken as a whole, rather than looking only at individual dyads or small sets of dyads. This can be done using readily-available and well-understood statistical techniques, although we suspect that specialized methods also will prove useful. The ability to analyze N-actor systems is particularly important once one begins to consider sub-state actors such as political factions and ethnic groups. Coding sub-state actors is relatively straightforward with automated coding and the techniques presented here should work as well with sub-state actors as they have with international actors. It would be interesting, for instance, to look for phase shifts in the Lebanon conflict using data that differentiates the activities of the various political groups within Lebanon.

The fact that we were generally unsuccessful in reducing the dimensionality of the system also argues for focusing on the analysis of N-actor systems rather than looking only at representative dyads or factor scores. In Schrodt and Gerner (1995), we analyzed the factor scores produced by our analysis using both discriminant and cluster analysis; the results were disappointing. In the discriminant analysis, the factor scores provide only 70% classification accuracy when we use the 21 scores with eigenvalues greater than 1.0, and only 59% accuracy when we use the six factors selected by a stepwise discriminant procedure. The clusters identified in the analysis of the factor scores are generally less stable and less plausible than those we found using the Goldstein scores.[16] The Levantine political subsystem is admittedly quite complex but in this respect it may be more typical of the situations that will be encountered in the post-Cold War world than were the bipolar and tripolar systems that have been the primary focus of most event data analyses in the past.

Second, as we noted above, the behaviors recorded in Reuters-based event data—at least in a well-covered area such as the Levant—are sufficiently detailed to discriminate correctly periods of time into the phases that were determined *ex post facto* by human analysts. Given the validity and reliability problems of event data, and the coding ambiguities and hindsight bias that affect human coding of phases, this conclusion was by no means guaranteed. To a more limited extent, the phases determined inductively by cluster analysis also correspond to the phases evident to a human analyst, although in this data set those phase assignments have much greater correspondence in the periods involving military and civil conflict than they do in periods involving negotiation. Because of this, we suspect that studies using crisis phase (whether human- or machine-coded) as an independent variable (for example, in assessing the effects of multilateral intervention on crisis outcome) will not be severely affected by specification error.

Finally, the next step in this analysis should be a study of the continuous movement of the N-actor behaviors with respect to the clusters, rather than the simple assignment of monthly data points to individual clusters. Because we cannot reduce the behavior of the system to two-dimensions, it is not operating with the simplicity of Figure 1; thus we need to use more sophisticated measures of cluster transitions. The cluster assignments in Figures 3a and 3b show two behaviors that need to be explored further: (a) the fluctuation in the cluster assignment prior to a phase transition; and (b) human-assigned phases such as *Taba*, *Madrid*, and *Oslo* that are characterized by rapid shifts between the statistically-assigned clusters. By focusing our analysis on the measurement of distances rather than the nominal assignment of phases, these characteristics might lead to some useful early warning indicators.

As we noted at the beginning of this article, the first wave of work in event-based statistical early warning—the DARPA-sponsored projects of the late 1970s and early 1980s—dwindled away without producing many useful results. This has led some analysts to conclude that efforts to develop such models are futile and that attention should be focused exclusively on qualitative early warning using the conventional techniques of intuitive political analysis.

In our opinion this assessment is premature and the past record of quantitative early warning models is not a prelude to the future. In several respects, the objectives of the DARPA-funded studies were beyond the capabilities of the technology available at the time. Dependent as they were on labor-intensive human-coding of events, most of the DARPA projects spent the bulk of their funds on data creation rather than analysis. Furthermore, days were required to conduct computer analyses on slow, batch-processing mainframes utilizing unfriendly statistical software. That same work can now be done in minutes. In short, the DARPA projects barely explored the domain of statistical techniques relevant to the early warning problem.

In the past decade and a half, great progress has been made in overcoming previous research limitations. Machine-coded event data are inexpensive, current, and, for a simple coding scheme such as WEIS, appear to be at least comparable, and quite possibly superior, to human-coded data. Global news services such as Reuters provide a far higher density of events than that provided by the older newspaper-based data sets, particularly in crisis-prone areas. A variety of techniques that were computationally infeasible or poorly understood in the 1970s are now routine and standard statistical methods such as regression, factor analysis, and discriminant analysis can be run in seconds rather than hours.

This is not to say that the problem of quantitative early warning is easy: After all, the record of human analysts using qualitative methods is hardly without flaw.[17] But with contemporary techniques the problem of early warning now appears to be considerably more tractable than it was in the era of the DARPA studies. These improved computational techniques must be coupled with a more sophisticated theoretical understanding of how crises escalate and are diffused; here, also, the models and concepts that are waiting to be systematically tested are much richer than those available two decades ago. As the utility of those various theories is evaluated through systematic tests using data from the post-Cold War period, quantitative early warning methods may prove to be a significant supplement to qualitative forecasting efforts.

Bibliography

Aldenderfer, Mark S. and Roger K. Blashfield. 1984. *Cluster analysis*. Newbury Park: Sage.

Alker, Hayward, Ted Robert Gurr, and Kumar Rupesinghe. 1995. Conflict early warning systems: An initial research program. Paper presented at the annual meeting of the International Studies Association, Chicago, February.

Andriole, Stephen J., and Gerald W. Hopple. 1984. The rise and fall of events data: From basic research to applied use in the U.S. Department of Defense. *International Interactions* 11:293-309.

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.

Bailey, Kenneth D. 1994. *Typologies and taxonomies: An introduction to classification techniques.* Thousand Oaks, CA: Sage Publications.

Bloomfield, Lincoln P., and Amelia C. Leiss. 1969. *Controlling small wars.* New York: Knopf.

Bloomfield, Lincoln P., and Allen Moulton. 1989. *CASCON III: Computer-aided system for analysis of local conflicts.* Cambridge: MIT Center for International Studies.

Bloomfield, Lincoln P., and Allen Moulton. 1997. *Managing international conflict.* New York: St. Martin's Press.

Boutros-Ghali, Boutros. 1994. Strengthening of the coordination of emergency humanitarian assistance of the United Nations. Secretary-General's Report to the General Assembly A/49/177, 2 September 1994.

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. 1976. *Managing interstate conflict, 1945-74: Data with synopses.* Pittsburgh: University of Pittsburgh University Center for International Studies.

Cahill, Kevin M., ed. 1996. *Preventive diplomacy: The therapeutics of mediation.* New York: Basic Books.

Choucri, Nazli, and Thomas W. Robinson, eds. 1979. *Forecasting in international relations: Theory, methods, problems, prospects.* San Francisco: W.H. Freeman.

Crocker, Chester, and Fen Osler Hampson, eds. 1996. *Managing global chaos: sources of and responses to international conflict.* Washington, D.C.: United States Institute of Peace Press.

Daly, Judith Ayres, and Stephen J. Andriole. 1980. The use of events/interaction research by the intelligence community. *Policy Sciences* 12:215-236.

Dedring, Jürgen. 1994. Early warning and the United Nations. *Journal of Ethno-Development* 4:98-105.

Dixon, William J. 1986. Reciprocity in United States-Soviet relations: Multiple symmetry or issue linkage. *American Journal of Political Science* 30:421-45.

Everitt, Brian. 1980. *Cluster analysis* (2nd ed.) New York: Wiley/Halsted.

Gaddis, John Lewis. 1992. *The United States and the end of the Cold War*. New York: Oxford University Press.

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* 38:91-119.

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.

Gurr, Ted R., and Barbara Harff. 1994. Conceptual, research and policy issues in early warning research: An overview. *Journal of Ethno-Development* 4,1:3-15.

Gurr, Ted R., and Barbara Harff. 1996. *Early warning of communal conflict and humanitarian crisis*. Tokyo: United Nations University Press, Monograph Series on Governance and Conflict Resolution.

Hopple, Gerald W. 1984. Computer-based early-warning: A staircase display option for international affairs crisis projection and monitoring. *National security crisis forecasting and management*, edited by Gerald W. Hopple, Stephen J. Andriole, and Amos Freedy, 47-84. Boulder: Westview Press.

Hopple, Gerald W., Stephen J. Andriole, and Amos Freedy, eds. 1984. *National security crisis forecasting and management*. Boulder: Westview Press.

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., and Jon C. Pevehouse. 1996. Potential validity problems in events data collection. *International Studies Notes* 21: 8-19.

Kim, Jae-On, and Charles W. Mueller. 1978. *Factor analysis: Statistical methods and practical issues*. Beverly Hills: Sage Publications.

King, Dennis. 1996. ReliefWeb: A new information management tool for the international humanitarian community. Paper presented at the conference on Risk Assessment and Early Warning Systems, University of Maryland at College Park, November.

Klecka, William R. 1980. *Discriminant analysis*. Beverly Hills: Sage Publications.

Laurance, Edward J. 1990. Events data and policy analysis. *Policy Sciences* 23:111-132.

Leatherman, Janie and Raimo Väyrynen. 1995. Structure, culture and territory: Three sets of early warning indicators. Paper presented at the annual meeting of the International Studies Association, Chicago, February.

Lund, Michael S. 1996. *Preventing violent conflicts: A strategy for preventive diplomacy.* Washington, D.C.: United States Institute for Peace Press.

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. 1993. Event data and DDIR. In *International event-data developments: DDIR phase II*, edited by Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes, 1-18. Ann Arbor: University of Michigan Press.

Mizuno, Jiro. 1995. Humanitarian early warning system: progress and prospects. United Nations: Department of Humanitarian Affairs.

Munton, Donald. 1978. *Measuring international behavior: Public sources, events and validity.* Dalhousie University: Centre for Foreign Policy Studies.

Norusis, Marija J. 1994. *SPSS professional statistics 6.1*. Chicago: SPSS, Inc.

Peterson, Sophia. 1975. Research on research: Events data studies, 1961-1972. In *Sage International Yearbook on Foreign Policy Studies* 3, edited by Patrick J. McGowan, 263-309. Beverly Hills: Sage Publications.

Rupesinghe, Kumar, and Michiko Kuroda, eds. 1992. *Early warning and conflict resolution.* New York: St. Martin's Press.

Schmeidl, Susanne. 1997. *From root cause to preventive diplomacy: Possibilities and limitations of the early warning of forced migration.* Westport, CT: Praeger.

Schrodt, Philip A. 1995. Event data in foreign policy analysis. In *Foreign policy analysis: continuity and change*, edited by Laura Neack, Jeanne A.K. Hey, and Patrick J. Haney, 145-166. New York: Prentice-Hall.

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. 1995. Empirical indicators of crisis phase in the Middle East, 1979-1995. Paper presented at the Twenty-Ninth North American meeting of the Peace Science Society, Columbus, Ohio, October.

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* 12:561-588.

Sherman, Frank L., and Laura Neack. 1993. Imagining the possibilities: The prospects of isolating the genome of international conflict from the SHERFACS dataset. In *International event-data developments: DDIR phase II,* edited by Richard L. Merritt, Robert G. Muncaster, and Dina A. Zinnes, 87-112. Ann Arbor: University of Michigan Press.

Singer, J. David, and Michael D. Wallace, eds. 1979. *To augur well: early warning indicators in world politics.* Beverly Hills: Sage.

Ward, Michael Don. 1982. Cooperation and conflict in foreign policy behavior. *International Studies Quarterly* 26:87-126.

Footnotes

[1] This is particularly true when one focuses on strategic political intelligence to the exclusion of the tactical military intelligence provided by satellite imagery and the monitoring of electronic communication.

[2] Sherman and Neack (1993) provide a review of the evolution of these data sets.

[3] The strengths and weaknesses of event data are discussed by Azar, Brody, and McClelland (1972), Burgess and Lawton (1972), Doran, Pendley, and Antunes (1973), Azar and Ben Dak (1975), Peterson (1975), Munton (1978), Goldstein and Freeman (1990), Daly and Andriole (1980), Merritt, Muncaster, and Zinnes (1993), Gerner et al. (1994), and Schrodt (1995).

[4] In other words, we converted individual X>Y events to a numerical score on the Goldstein scale, then totaled these numerical scores by month to obtain a net cooperation score for each directed dyad over time. Schrodt and Gerner (1994) gives several time series plots of the data for the 1982-1993 period.

The high dimensionality of this space makes it difficult to visualize. This is not a new problem in event data analysis: The response of most of the earlier event data studies was to ignore N-actor systems and instead either focus on a small number of dyads (e.g., Ward 1982; Dixon 1986; Goldstein and Freeman 1990) or examine the interactions of one actor with a number of states (e.g., Howell and Barnes 1993). In situations where there are a small number of clearly dominant dyads—for example, the USA-USSR-PRC triad of the Cold War—this is effective. Our research suggests, however, that in a complex system such as the Middle East, determining *a priori* the relative importance of various dyads is difficult and may distort the analysis.

[5] NEXIS is searched using keywords that can be arranged into Boolean statements. The source texts for this data set were located with the search command (ISRAEL! OR JORDAN! OR EGYPT! OR LEBAN! OR SYRIA! OR PLO OR PALEST!). The "!" is a wild card character that matches any word beginning with the preceding letters; so "PALEST!" picks up not only "Palestinian," for example, but "Palestinians" and "Palestine." The initial download generated about 100,000 events. Many of these are outside the 54 directed dyads considered in this study, however; the smaller data set contains about 40,000 events.

[6] We have excluded the USA>USR and USR>USA dyads from our analysis since most of their interactions did not deal with the Middle East. Inclusion of these dyads in the data set makes relatively little difference in the analysis. For instance, the same five factors are found in the factor analysis and the variance explained by each factor changes by less than ±0.2%: Both of the dyads have their maximum loading on the "Camp David" factor with a loading of 0.336 for USA>USR and 0.505 for USR>USA. The accuracy of the discriminant analysis increases from 90.1% to 91.1% when the two superpower dyads are included.

[7] The data and KEDS coding dictionaries are also being archived in the ICPSR's "Publications-Related Archive."

[8] In other words, varimax rotation tries to associate each variable with one and only one factor. All analyses were done using SPSS 6.1 for the Macintosh Power PC.

[9] In the unrotated solution, the first two factors also emphasize the Lebanon and Israeli-Palestinian disputes but with less separation of the variables.

[10] Alternatively, accurate phase classification based on event data might be possible through a nonlinear method such as an expert system, neural network, or genetic algorithm, but not through discriminant analysis.

[11] The Wilks' lambda reported for each dyad is the proportion of the discrimination between the phases that is *not* explained by the dyads up to and including that dyad. For example, the ISR>PAL, LEB>ISR, PAL>UAR and ISR>JOR dyads together account for 75% of the discrimination between phases that is done by the model.

[12] We also allowed this assignment process to iterate but this had little effect on the final cluster centers.

[13] Moving $\text{Mode}_t = \text{Mode}(X_t, X_{t+1}, X_{t+2}, X_{t+3})$ Ties were set to X_t if the tie was generated by the pattern XXYY and set to $\text{Mode}(X_{t-1}, \ldots, X_{t+4})$ for any other tied pattern. Cluster "X" contains those points where the expanded mode also produced a tie, where no mode could be computed because the 4-month interval contained four different clusters, or where the mode occurred in the pattern YZXX, i.e. the actual transition did not occur until t+2. Neither of these assumptions are critical to the smoothing. Graphs containing the unsmoothed cluster assignments can be found at the URL http://www.ukans.edu/~keds/jcr97 or obtained from the authors.

[14] The only K-Means analysis we have done that successfully finds a "peace process" cluster used the 21 factor scores that had eigenvalues greater than 1.0. In that experiment, the *Camp David* period and the *Madrid-Oslo* period were each in distinct clusters (without a transition between *Madrid* and *Oslo*), but the intervening period did not show plausible patterns—in fact, the factor score clustering was the only analysis not to show any part of the *intifada* as a distinct phase.

[15] Table 4 reports the first three factors; the analysis found 7 factors with eigenvalues greater than 1.0 of the ISR>LEB case and 8 such factors for the ISR>PAL case.

[16] Part of the problem with generating political plausible interpretations of the factor scores may be that the factors are orthogonal (uncorrelated): This is useful for statistical purposes such as regression but makes the factors very difficult to interpret politically because political interactions are generally *not* orthogonal—events in Lebanon continually affected Israeli-Palestinian interactions and vice versa.

[17] See Gaddis (1992) for an unusually thorough assessment of this issue.

Biographical Statements

Philip A. Schrodt

Philip A. Schrodt is a professor of political science at the University of Kansas. His major areas of research are formal models of political behavior, with an emphasis on international politics, and political methodology. He has published articles on these topics in the *American Journal of Political Science, Journal of Conflict Resolution, International Studies Quarterly* and *Political Analysis*, and is the author of the Kansas Event Data System (KEDS) automated coding program.

Deborah J. Gerner

Deborah J. Gerner is an associate professor of political science at the University of Kansas. She is the editor of Understanding the Contemporary Middle East (1998) and author of One Land, Two Peoples: The Conflict over Palestine (1990, 1994). Her research on Arab-Israeli relations, U.S. foreign policy and the analysis of event data has been published in more than a dozen books and, *inter alia*, The American Journal of Political Science, Arab Studies Quarterly, International Studies Quarterly, and The Journal of Arab Affairs. Gerner travels regularly to the Middle East and in 1996 spent six months on a Fulbright grant teaching international studies at Birzeit University (West Bank).

TABLE 1

A Priori Determination of Behavioral Phases in the Levant Region, 1979-1995

Label	Dates	Months	Defining Characteristic
Camp David	Jun.79-May.82	35	Before Israel's 1982 invasion of Lebanon
Lebanon	Jun.82-May.85	36	Israeli troops in Lebanon
Taba	Jun.85-Nov.87	30	Israeli withdrawal to south of the Litani until *intifada*
intifada	Dec.87-Jul.90	32	Palestinian *intifada*
Kuwait	Aug.90-Oct.91	15	Iraq's invasion of Kuwait until start of Madrid talks
Madrid	Nov.91-Aug.93	22	Bilateral and multilateral peace talks
Oslo	Sept.93-Jun.95	22	Oslo peace process

TABLE 2

Factor Analysis of the Monthly Aggregated Goldstein Scores

FACTOR 1
Lebanon Conflict: Regional

ISR>SYR	.77	-.04	.11	-.02	.08
ISR>LEB	.68	-.19	.27	-.11	-.10
LEB>ISR	.62	-.25	.25	.06	-.07
SYR>ISR	.61	.07	-.12	.07	.01
SYR>LEB	.61	.11	.04	-.05	.02
LEB>PAL	.56	-.01	-.08	.22	-.10
LEB>SYR	.47	.12	.05	.07	.31
PAL>LEB	.44	.20	-.10	.22	-.20
USR>ISR	.42	.02	-.08	.09	-.02
SYR>PAL	.22	.12	-.01	.03	-.21
USR>UAR	-.21	-.07	.04	.11	-.04

FACTOR 2
Israeli-Palestinian Conflict

PAL>USA	-.01	.67	.17	.06	.02
USA>PAL	-.07	.67	.40	.11	-.01
PAL>ISR	.17	.62	.33	-.21	-.02
JOR>ISR	-.03	.54	.00	.00	-.04
ISR>JOR	.07	.47	-.23	-.00	.08
USA>ISR	.08	.46	.24	.10	-.05
USA>JOR	.05	.45	.08	.03	-.00
ISR>PAL	.15	.43	.44	-.20	-.18
PAL>JOR	.09	.29	-.14	-.04	.02
JOR>USA	-.12	.28	-.17	.03	-.18
ISR>USA	.01	.27	.31	.01	-.05

FACTOR 3
Camp David

USA>UAR	.03	.00	.66	-.07	-.03
UAR>ISR	-.00	.06	.52	-.09	-.01
UAR>USA	-.23	-.04	.51	.02	-.18
USR>LEB	.13	.13	.50	.08	.09
ISR>PAL	.15	.43	.44	-.20	-.18
LEB>USR	-.08	.05	.41	.41	.08
ISR>UAR	-.00	.12	.37	.06	.08
ISR>USA	.01	.27	.31	.01	-.05
JOR>PAL	-.02	.15	-.22	-.07	.08

FACTOR 4
Lebanon: USA

USA>LEB	.09	-.18	.13	.76	.11
LEB>USA	.05	-.19	.19	.65	.08
USA>SYR	.10	.18	-.13	.64	.03
SYR>USA	.24	.21	-.27	.56	-.01
PAL>SYR	.04	.21	-.02	.41	-.02
LEB>USR	-.08	.05	.41	.42	.08
UAR>JOR	.07	.10	-.18	-.39	.07
USR>SYR	.07	.04	.10	-.26	.22

FACTOR 5
Lebanon: Jordan

LEB>JOR	-.00	-.04	.08	.07	.72
SYR>JOR	-.04	-.04	.03	-.05	.64
JOR>LEB	-.03	-.22	-.01	.12	.63
JOR>SYR	-.03	.06	-.14	-.07	.60
UAR>LEB	.03	-.07	.19	.01	-.45

No loading above 0.20 on first five factors

USR>PAL	-.02	-.08	.10	-.01	.03
ISR>USR	.17	.08	-.04	.13	.12
JOR>UAR	.01	.09	-.15	-.09	-.00
JOR>USR	-.19	-.10	-.02	.14	.07
LEB>UAR	-.03	-.01	-.05	-.04	-.09
PAL>UAR	.04	.20	-.04	.06	.06
PAL>USR	.14	-.11	.08	.03	.13
SYR>UAR	.12	-.02	-.01	.07	-.01
SYR>USR	-.05	-.00	.11	-.04	.00
UAR>PAL	-.01	.15	.20	-.12	.03
UAR>SYR	.02	-.15	-.07	-.04	-.05
UAR>USR	-.15	-.07	.05	.13	-.15
USR>JOR	-.01	-.14	.07	.09	.02

Variance Explained by Factors:

F1:	7.7%
F2:	6.3%
F3:	4.9%
F4:	4.6%
F5:	4.3%

SOURCE: Computation by the authors. Factor analysis of all directed dyads; varimax rotation. Variables are sorted by maximum loading; they are included in two factors if the difference between the two highest loadings is <0.05. Variables are included in a factor if the loading is >0.20 in absolute value.

TABLE 3a.
Discriminant Analysis using Monthly Aggregated Goldstein Scores for All Dyads to Predict A Priori Phase

	Predicted							
Actual	Camp David	Lebanon	Taba	Intifada	Kuwait	Madrid	Oslo	N
Camp David	**100.0%**	0%	0%	0%	.0%	0%	0%	35
Lebanon	2.8%	**86.1%**	5.6%	2.8%	0%	2.8%	0%	36
Taba	0%	0%	**86.7%**	0%	3.3%	10.0%	0%	30
Intifada	3.1%	0%	3.1%	**90.6%**	0%	3.1%	0%	32
Kuwait	0%	0%	0%	0%	**100.0%**	0%	0%	15
Madrid	4.5%	0%	4.5%	0%	.0%	**77.3%**	13.6%	22
Oslo	0%	0%	0%	0%	0%	9.1%	**90.9%**	22

Percent of cases correctly classified: 90.10%

Discriminant Function	Var Explained	Cumulative Pct	Wilks' Lambda	Signif
1	28.74	28.74	.024	<.001
2	24.44	53.18	.072	<.001
3	20.62	73.80	.196	<.001
4	11.12	84.92	.378	<.001
5	8.61	93.53	.650	.030

TABLE 3b.
Stepwise Discriminant Analysis using Monthly Aggregated Goldstein Scores to Predict A Priori Phase

	Predicted							
Actual	Camp David	Lebanon	Taba	Intifada	Kuwait	Madrid	Oslo	N
Camp David	**85.7%**	2.9%	8.6%	0%	2.9%	0%	0%	35
Lebanon	2.8%	**63.9%**	16.7%	2.8%	5.6%	8.3%	0%	36
Taba	10.0%	6.7%	**56.7%**	10.0%	3.3%	13.3%	0%	30
Intifada	0%	0%	9.4%	**68.8%**	12.5%	6.3%	3.1%	32
Kuwait	6.7%	0%	6.7%	0%	**73.3%**	13.3%	0%	15
Madrid	9.1%	4.5%	4.5%	9.1%	0%	**68.2%**	4.5%	22
Oslo	0%	0%	4.5%	0%	4.5%	13.6%	**77.3%**	22

Percent of "grouped" cases correctly classified: 70.31%

Discriminant Function	Var Explained	Cumulative Pct	Wilks' Lambda	Signif
1	32.95	32.95	.156	<.001
2	27.45	60.41	.311	<.001
3	19.97	80.38	.535	<.001
4	10.89	91.27	.746	<.001
5	5.54	96.80	.896	.006

TABLE 4.

Factor Analysis of the Monthly Event Counts in the 2-Digit WEIS Categories

	ISR>LEB			ISR>PAL		
Factor	Eigenvalue	Variance Explained	Corr*	Eigenvalue	Variance Explained	Corr*
1	7.71	35.1%	-.71	5.59	25.4%	-.80
2	1.72	7.9%	.45	2.11	9.6%	.12
3	1.35	6.1%	.13	1.71	7.8%	-.12

* "Corr" is the bivariate correlation of the factor with the Goldstein score.

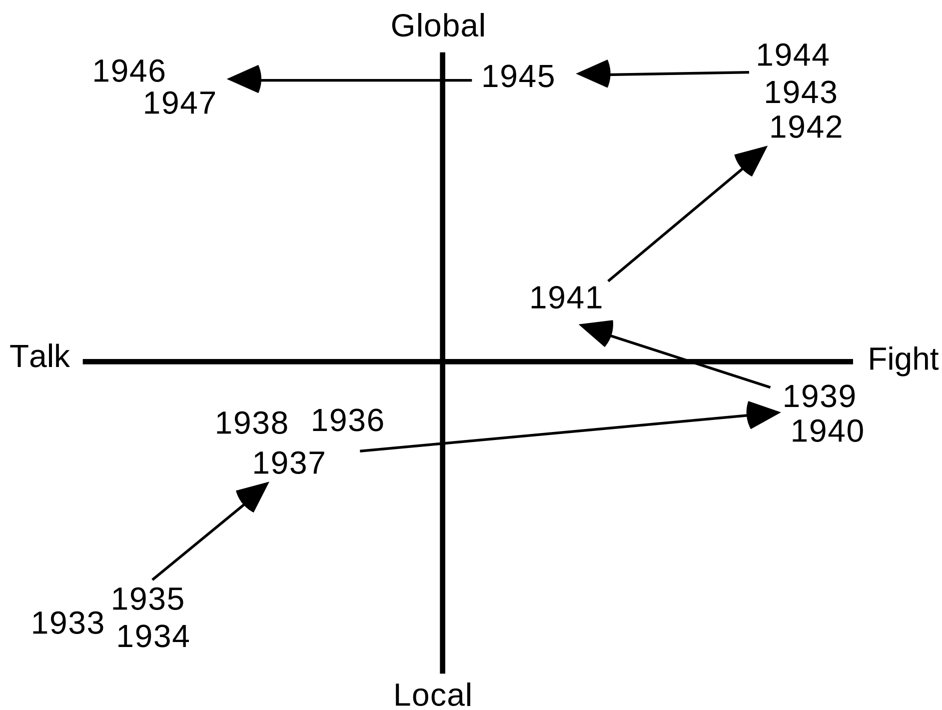


Figure 1: Hypothetical Representation of Phases during the WWII Period

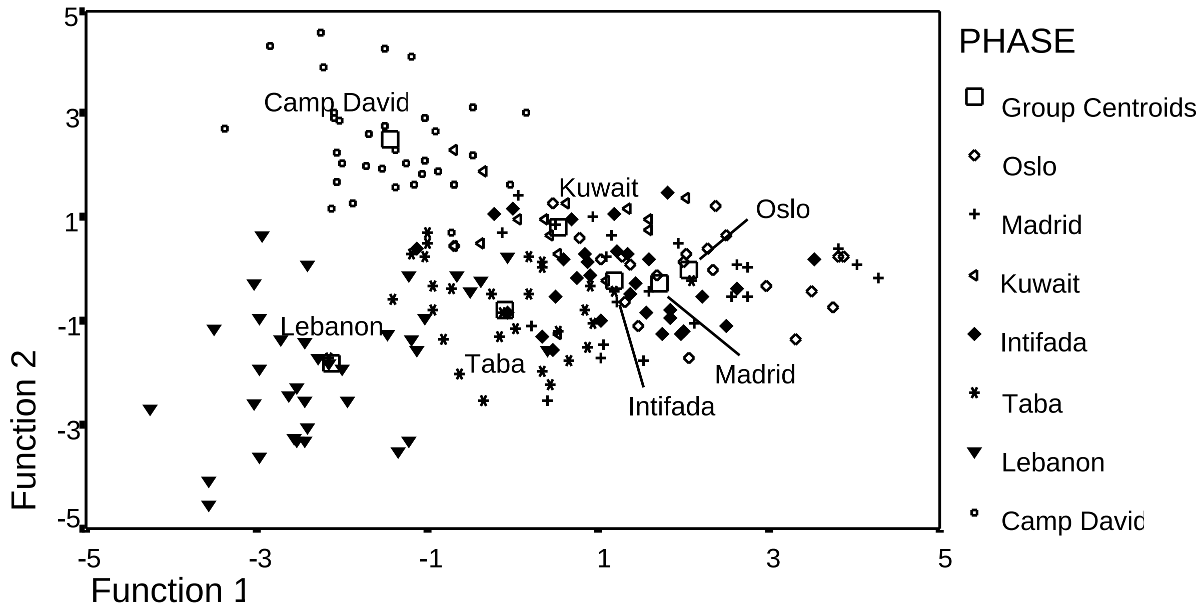
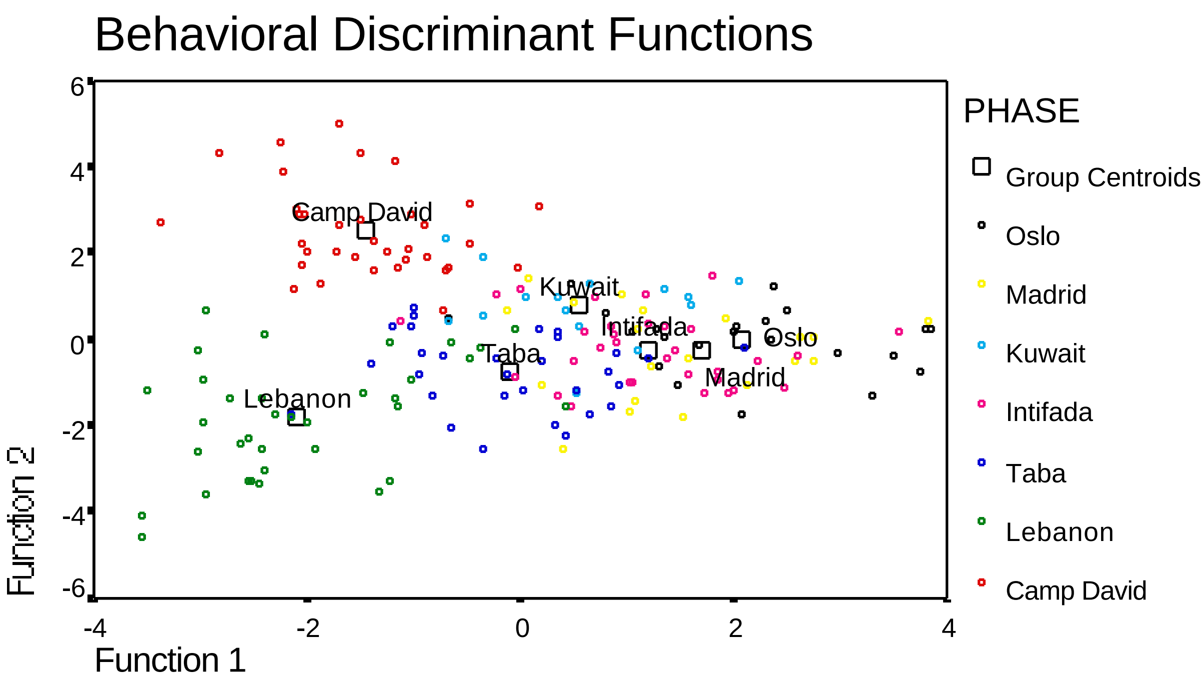


Figure 2: Discriminant Space—Functions 1 and 2



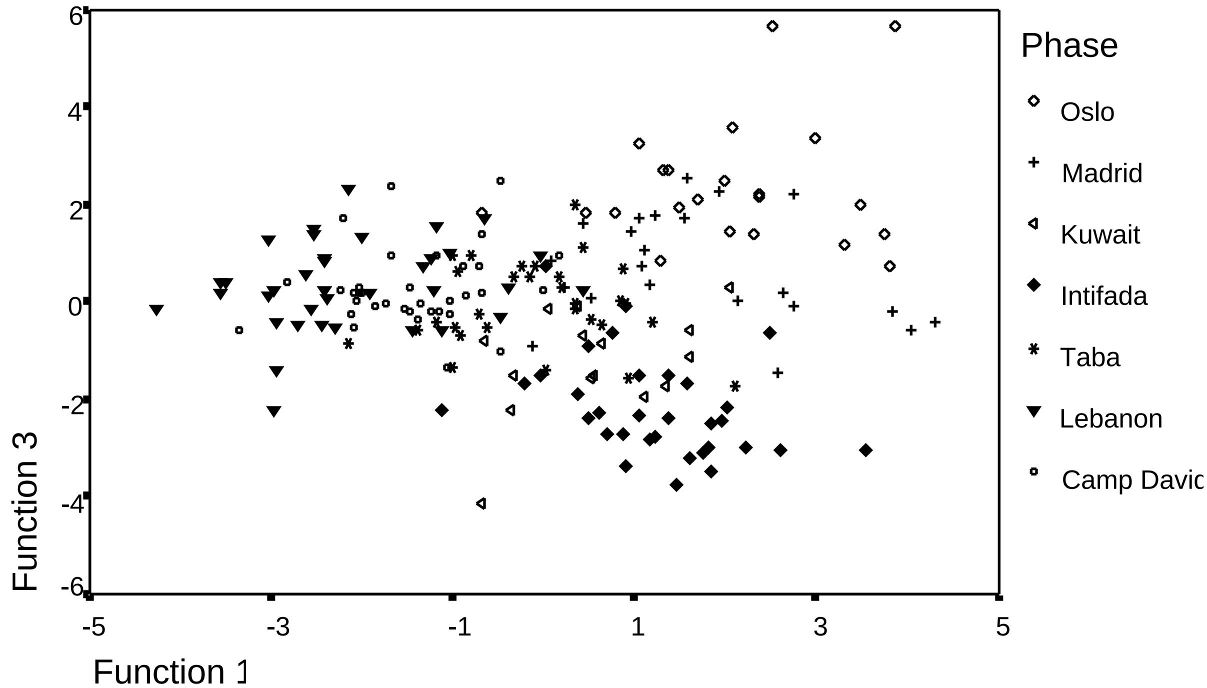
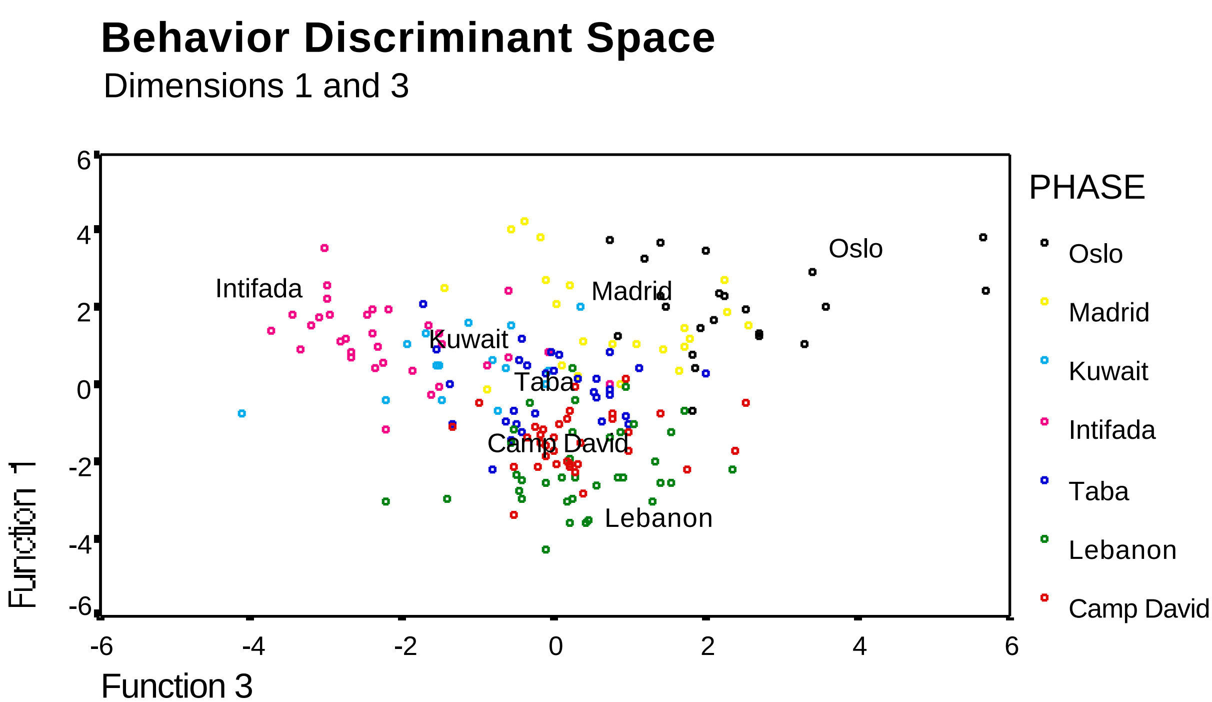
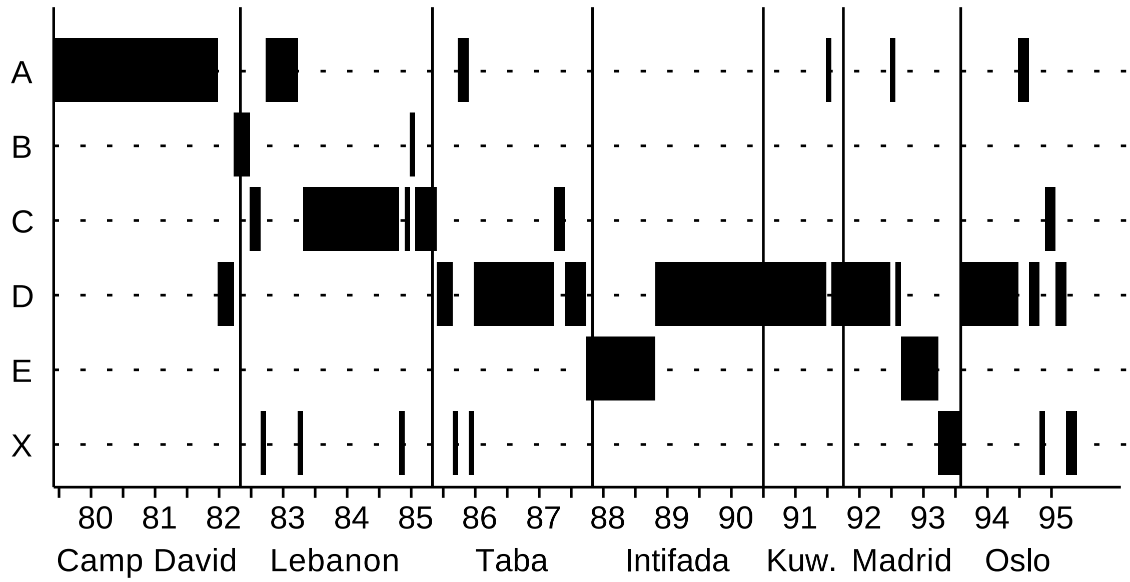


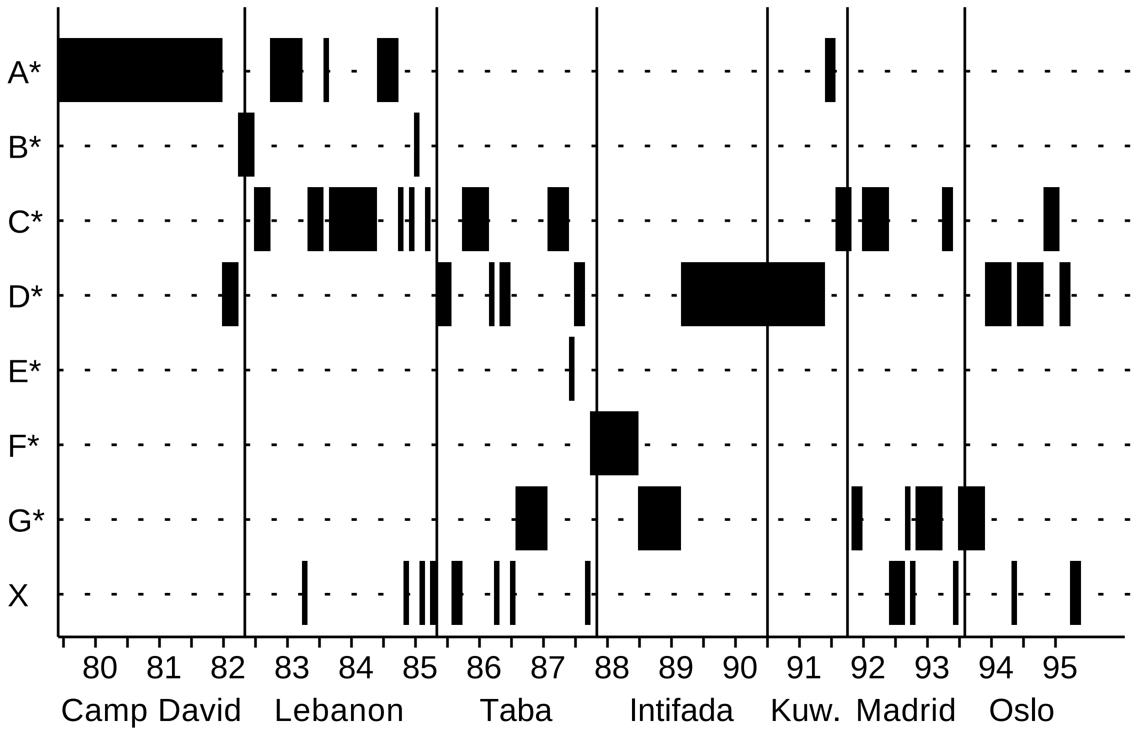
Figure 3: Discriminant Space—Functions 1 and 3





Note: Months in the "X" cluster are those for which a mode could not be computed.

Figure 3: K-Means Clusters for Monthly Aggregated Goldstein Scores for All Dyads, with Modal Smoothing



Note: Months in the "X" cluster are those for which a mode could not be computed.

Figure 4: K-Means Clusters for Monthly Aggregated Goldstein Scores for Israeli-Palestinian Conflict Dyads Only, with Modal Smoothing