The Creation of CAMEO (Conflict and Mediation Event Observations): An Event Data Framework for a Post Cold War World

DEBORAH J. GERNER, PHILIP A. SCHRODT, OMUR YILMAZ

Department of Political Science, University of Kansas

RAJAA ABU-JABR

Independent researcher, Nashville, Tennessee

Prepared for delivery at the 2002 Annual Meeting of the American Political Science Association, 29 August – 1 September 2002.

This research was supported in part by the U.S. National Science Foundation grant SES-0096086 and by the General Research Fund of the University of Kansas. Rebecca Haas, Bradley Lewis, Rebekah Moses, P. M. Picucci, Joseph Pull, Almas Sayeed, Erin Simpson, and Nooshin Soltani assisted in the development of the CAMEO coding framework and the elaboration of the coding dictionaries. The CAMEO codebook, all data sets, dictionaries, and software discussed in this paper can be downloaded from the KEDS project web site: http://www.ku.edu/~keds. An earlier version of this paper was presented at the annual meeting of the International Studies Association, March 2002, New Orleans.

Abstract

The Conflict and Mediation Event Observations (CAMEO) framework is a new event data coding scheme optimized for the study of third party mediation in international disputes. The World Events Interaction Survey (WEIS) framework that the authors used in previous event data research has a number of shortcomings, including vagueness in and overlap of some categories, and a limited applicability to contemporary issues involving non-state actors. The authors have addressed these and other problems in constructing CAMEO and have produced far more complete documentation than is available for WEIS.

CAMEO has been developed and implemented using the TABARI automated coding program and has been used to generate data sets for the Balkans (1989-2002; N=71,081), Levant (1979-2002; N=139,376), and West Africa (1989-2002; N=18,519) from Reuters and Agence France Presse reports. This article reports statistical comparisons of CAMEO-coded and WEIS-coded data for these three geographical regions. CAMEO and WEIS show similar irregularities in the distribution of events by category. In addition, when the data are aggregated to a general behavioral level (that is, into verbal cooperation, material cooperation, verbal conflict and material conflict), most of the data sets show a high correlation ($r > 0.90$) in the number of WEIS and CAMEO events coded per month. Finally, there is a significant correlation ($r > 0.57$) between the count of CAMEO events specifically dealing with mediation and negotiation, and a pattern-based measure of mediation the authors developed earlier from WEIS data. CAMEO thus appears to maintain coverage of events typically coded by WEIS while adding enhanced precision and stronger coverage of additional activities such as mediation that are of increasing scholarly interest in the twenty-first century.

Introduction

The Kansas Event Data System (KEDS) project develops automated natural language processing software, creates specialized event data sets on international political behavior, and analyzes these data statistically.[1] Our initial machine coding system, KEDS, was validated against both the textual record and human-coded events (Gerner et al., 1994; Schrodt & Gerner, 1994) and has been used by scholars looking at interactions in Northern Ireland (Thomas, 1999), the Balkans (Goldstein & Pevehouse, 1997; Pevehouse & Goldstein, 1999; Schrodt & Gerner, 2001; Schrodt et al., 2001), the Middle East (Gerner & Schrodt, 1998; Schrodt, 1999; Schrodt & Gerner, 2000; Rasler, 2000; Gerner et al., 2001; Goldstein et al., 2001), West Africa (Huxtable, 1997), and the United States (Wood & Peake, 1998). In 2000, Schrodt created a new program, Textual Analysis by Augmented Replacement Instructions (TABARI), as the successor to the KEDS software.[2] We have recently begun to use TABARI to code events relevant to third party mediation in three geographical regions: the Middle East (1979-2002), the Balkans (1989-2002), and West Africa (1989-2002).

This paper describes a new event data coding system—Conflict and Mediation Event Observations (CAMEO)—that we have developed for this project and provides some statistical comparisons between the new framework and the World Event Interaction Survey (WEIS) codes with which we have been working for a number of years. We first discuss the factors that motivated our creation of CAMEO, including consideration of some of the problems we found with the WEIS framework. We then outline the intellectual and mechanical processes involved in constructing CAMEO and describe some of its distinguishing features. Third, we compare our CAMEO and WEIS data sets in terms of the numbers and types of events coded as well as looking specifically at how each measures mediation activities. Our findings suggest that CAMEO maintains excellent coverage of events typically coded by WEIS while adding increased precision and stronger coverage of additional activities such as mediation that are of increasing scholarly interest in the twenty-first century. (The Appendix summarizes the basic CAMEO and WEIS frameworks.)

Why a New Coding Framework?

For several decades, event data research has been dominated by two coding frameworks: Charles McClelland's WEIS (1976) and the Conflict and Peace Data Bank (COPDAB) developed by Edward Azar (1982). Both were created during the Cold War and assumed a 'Westphalian-Clausewitzian' political world in which sovereign states react to each other primarily through official diplomacy and military threats. While innovative when first created, these coding systems are not optimal for dealing with contemporary issues such as ethnic conflict, low-intensity violence, organized criminal activity, and multilateral intervention. McClelland (1983: 177) himself viewed WEIS as only a 'first phase'; he certainly did not anticipate that it would continue to be used, with only minor modifications, for four decades.

---

[1] Event data—nominal or ordinal codes recording the interactions between political actors as reported in the open press—break down complex activities into a sequence of basic building blocks that can be analyzed statistically.

[2] Like KEDS, TABARI—which is 'open-source' code and available for the Linux, Macintosh, and Windows operating systems—uses a computational method called 'sparse parsing.' Instead of trying to decipher a sentence fully, TABARI determines only the parts required for event coding—for instance, political actors, compound nouns and compound verb phrases, and the references of pronouns—and then employs a large set of verb patterns to determine the appropriate event code. Sparse parsing techniques can be used successfully on unedited news wire text such as lead sentences from the Reuters and Agence France Presse news services. Automated event data coding is more reliable and transparent than human coding and—once the actor and verb dictionaries have been developed—automated coding is about seven-million times faster than human coding. See Gerner et al (1994), Schrodt & Gerner (1994), Schrodt, Davis & Weddle (1994), Bond et al. (1997), Thomas (1999), and King & Lowe (2001) for additional discussions of automated coding.

Event categories present in WEIS and COPDAB have both conceptual and practical shortcomings. For instance, WEIS has only a single cue category of *Military engagement* that must encompass everything from a shot fired at a border patrol to the strategic bombing of cities. (The phrase 'cue category' refers to the broad two-digit codes, as opposed to the more specific three and four digit subcategories.) COPDAB contains just 16 event categories, spanning a single conflict-cooperation continuum that many researchers consider inappropriate. Although there have been efforts to create alternative coding systems—most notably Leng's (1987) Behavioral Correlates of War (BCOW)—WEIS and COPDAB remain the predominant frameworks in the published literature. The 'lock-in' of these early coding systems is readily explained by the time consuming nature of human event coding from paper and microfilm sources. Because human coders typically produce between five and ten events per hour, and a large data set contains tens of thousands of events, experimental recoding was simply not feasible. Established protocols for training and maintaining consistency among coders further constrained efforts to improve WEIS and COPDAB once these were institutionalized. As a consequence, endeavors such as Tomlinson's (1993) modification of WEIS and the Global Event Data System (GEDS) project extensions of COPDAB (Davies & McDaniel, 1993) produced only marginal changes.

Automated coding, in contrast, allows researchers to experiment with alternative coding rules that reflect a particular theoretical perspective or interest in a specific set of issues. The effort involved in implementing a new or modified coding system, once it has been developed, is relatively small because most of the work can be done within the dictionary of verb phrases. In most cases verb phrases can be unambiguously assigned to appropriate new categories, while obscure phrases are either removed or modified. This elimination of questionable phrases itself represents an improvement in the coding system. Even a long series of texts spanning multiple decades can then be recoded in a few minutes. This allows researchers to focus on maximizing the validity of a particular coding scheme since the automated coding process itself guarantees the reliability of the system. New coding frameworks that have use automated coding include Bond et al.'s (1997) Protocol for the Analysis of Nonviolent Direct Action (PANDA) and more recently the Integrated Data for Event Analysis (IDEA) system (http://vranet.com/idea/; also see King & Lowe, 2001).

In the early stages of the KEDS project, when we were initially developing automated coding programs, we felt it was important to work with an existing framework so that we could directly compare human-coded and machine-coded data (Schrodt & Gerner, 1994). For a variety of reasons, we selected WEIS which, despite some obvious drawbacks, was "good enough" for our initial analyses. However, we recently decided to abandon WEIS. Several considerations motivated this choice. First and foremost was our long-standing concern regarding numerous ambiguities, overlaps, and gaps within the WEIS framework. In addition, the distribution of events in WEIS is quite irregular and several of the 2-digit cue categories generate almost no events; we thought that perhaps we could improve on this. Third, we wanted to eliminate distinctions among actions that, while analytically discrete, could not be consistently and reliably differentiated from the existing news source materials. Finally, as indicated above, the Cold War perspective that permeates WEIS makes it an inappropriate tool for studying contemporary international interactions. Consequently, we developed CAMEO, which is specifically designed to code events relevant to the mediation of violent conflict.

Specific Problems with WEIS

Our coding experience with WEIS has led us to recognize major drawbacks and weaknesses in several of the WEIS categories. There are a number of events that do not appear to fit anywhere, instances where different types of events are placed in a single category, and categories that appear to us to be virtually identical. Each of these problems raises issues of validity and reliability. It is difficult to maintain conceptual consistency within categories when the definitions are broad, vague, or unclear. Problems encountered with WEIS are exacerbated due to the lack of a fully specified standard codebook. We based our development of coding dictionaries on the version of the WEIS codebook available through the Inter-university Consortium for Political and Social Research (ICPSR)

(McClelland, 1976). The section of the codebook dealing with event categories is quite short—about five pages—and provides only limited guidance. Since McClelland never intended that WEIS would become a *de facto* coding standard, the ICPSR WEIS codebook was meant primarily to be a proof-of-concept. More extensive WEIS-based codebooks—Sherwin and VanBeers (1976), Third Point Systems (1985), and Tomlinson (1993)—are not easily available and their proposed extensions have not been widely disseminated and accepted.

A discussion of cue category *Force* (WEIS 22) illustrates some of WEIS's shortcomings. Most markedly, the *Force* subcategories combine a myriad of events that range from violent civilian demonstrations to military occupations of states. These event types are substantially different in their levels of intensity, the types of actors typically involved, the means employed, and the ends pursued. Combining all of these dissimilar events into a single cue category creates analytical confusion. This is a crucial concern since scholars in their analysis typically use data aggregated at the two-digit cue category level.

The broadness of the subcategories under *Force* adds to its weakness. As already mentioned, *Military engagement* (WEIS 223), for instance, encompasses all types of violent activities involving military forces, such as the use of small arms, artillery and tank attacks, aerial bombings, closures, incursions, and occupations. A shooting incident between military forces and a group of stone-throwing civilians is distinguished neither from conventional acts of war between troops nor from uses of massive unconventional force, such as chemical, biological, radiation, or nuclear attacks. Vagueness within this subcategory is bolstered by its potential overlap with event codes under different categories. For example, according to the ICPSR WEIS codebook, *Seize position or possession* (WEIS 211) 'may also be used when a nation militarily takes or occupies another's territory,' even though these activities clearly constitute acts of *Military engagement* at the same time. This leaves coders unable to make consistent and objective decisions, thereby begetting reservations about the reliability of verb dictionaries and the final data sets. Questions of validity also arise: It isn't clear what types of events we seek to measure by each event form and whether those match what we end up capturing.

Nonmilitary injury-destruction (WEIS 222) is similarly problematic in its aggregation of substantially different event types ranging from demolition of houses to stone throwing, suicide bombings, kidnappings, assassinations, and guerilla attacks. In other words, capturing of a village by the Congolese Rally for Democracy after heavy fighting with state troops, kidnapping of tourists by the Abu Sayyaf, stone throwing by protesters at a demonstration, and the September 11 attacks against New York city and Washington, D.C. are all assigned the same code. Furthermore, *Nonmilitary demonstration* (WEIS 181), which includes activities such as 'marching, picketing, stoning', clearly overlaps with *Nonmilitary injury-destruction* in cases of demonstrations that become violent in some way. This overlap again raises the possibility of inconsistent coding unless systematic distinctions are provided in an elaborate codebook and the coders are trained correspondingly.

A *theoretical* (as opposed to practical) problem with the *Force* subcategories is that *Military engagement* is distinguished from *Nonmilitary injury-destruction* only in terms of who the actor is rather than the nature of the event. In other words, the distinction that WEIS makes between the two codes is a distinction between acts initiated by state forces and those initiated by all other actors (ranging from equipped and organized guerilla groups to individual dissenters). While it is essential to distinguish violent acts by organized armed forces (whether they belong to an internationally recognized state or not) from those by anomic groups of protesting individuals, it is troublesome to make the same distinction between acts of very similar characteristics—artillery attack by a guerilla group and the same type of assault by a state-sponsored group—based solely on the political identity of the actor.

Another area of confusion we encountered with WEIS results from cue categories that, while conceptually discrete, are practically impossible to distinguish from each other given the nature of

news leads. In particular, we have had significant difficulties systematically differentiating *Promise* (WEIS 05) and *Agree* (WEIS 08), *Grant* (WEIS 06) and *Reward* (WEIS 07), and *Request* (WEIS 09) and *Propose* (WEIS 10). For instance, even though the words 'promise' and 'agree' clearly hold different meanings, news leads often contain other verbs that indicate some form of future commitment and cannot be easily or consistently categorized as either a promise or an agreement. The distinction is even harder to maintain at the level of subcategories: *Agree* and *Request*, for instance, include subcategories that are practically impossible to demarcate: *Ask for policy assistance* (WEIS 092) and *Urge or suggest action or policy* (WEIS 102). Does 'calling for support' constitute an act of asking for assistance or one of urging a certain policy? Is 'urging' substantially and measurably different from 'asking' as to warrant its own separate code? Our experience with WEIS indicates that it is not. The same applies for the distinction between *Urge or suggest action or policy* and *Request action/call for* (WEIS 094).

Finally, there are the potential overlaps, and the difficulties of conceptual and practical differentiation, with respect to *Demand* (WEIS 15), *Warn* (WEIS 16), and *Threaten* (WEIS 17). Although in some instances these can mean clearly different things, in other uses the distinction among them becomes quite subjective. For instance, when US President George W. Bush says 'Palestinians must stop suicide bombings', is this a demand, a warning, or a threat? The answer would depend not only on the context but also on the worldview and the political perspective of each respondent. The ICPSR codebook description for the only subcategory of *Warn—Give warning*—illustrates, rather than solves, the problem: 'Occasionally the words "demand" or "threaten" are used in news items which should be coded as warnings'. In other words, in some cases even phrases with the verbs 'demand' and 'threaten' are to be coded under *Warn*, rather than the respective categories of these verbs. It is not any less ambiguous when other verbs are used to imply demands, warnings, or threats. The human coder is expected to make a judgment call, leading to inconsistent coding; in automated coding, a given verb phrase can only be coded a single way but this means simply picking one code or the other while knowing that at least some of the time the code will be wrong.

In summary, WEIS contains a variety of significant weaknesses that cannot be corrected without making substantial changes in the framework. Despite the earlier commitment we made to WEIS, and despite the body of existing research based on WEIS, we concluded that it is time to make those changes. The result is CAMEO.

The CAMEO Framework

Following the lead of IDEA, we initially intended CAMEO to be an extension of WEIS. The first phase of CAMEO's development involved the addition of cue and subcategories that we found theoretically necessary for the study of mediation and conflict, while keeping most of the WEIS cue categories intact. The next phase involved looking for sample leads and writing definitions for the codebook. This process led to the realization that some of the distinctions we wanted to make for theoretical reasons were simply not possible given the nature of the news leads. For instance, as we have indicated above, *Promise* (WEIS 07) is almost indistinguishable from *Agree* (WEIS 08) unless the word 'promise' is used in the lead. Therefore, we eventually ended up merging the two into a single cue category—*Agree* (CAMEO 06)—that includes codes representing all forms of future commitment. Similarly, because verbs such as 'call for', 'ask for', 'propose', 'appeal', 'petition', 'suggest', 'offer', and 'urge' are used interchangeably in news leads to refer to closely related activities, we combined *Request* and *Propose* into a single cue category—*Request/Propose* (CAMEO 05). We made similar decisions with respect to other WEIS categories such as *Grant* and *Reward*, and *Warn* and *Threaten*.

While developing CAMEO, we paid significant attention to achieving consistency in our additions to and/or modifications of original WEIS categories in order to ensure the creation of a

conceptually coherent and complete coding scheme. In other words, having the cue category of *Approve* (CAMEO 03) necessitated the addition of *Disapprove* (CAMEO 11), which incorporated *Accuse* (WEIS 12) and our new addition *Official Protest* (CAMEO 113). Maintaining the cue category of *Reduce Relations* from WEIS, albeit in a modified fashion, directed us to create a parallel category that captures improvements in relations—*Cooperate* (CAMEO 04). Even though they might not be represented by exact antonyms, we tried to make sure that conceptual opposites of each cue and subcategory exists within the coding scheme.

In addition, we made CAMEO consistent with respect to the order of its main cue categories. Unlike WEIS and IDEA, we started with the most neutral events and moved gradually from cooperation to conflict categories. While the initial coding category in WEIS and IDEA is *Yield*, CAMEO starts with *Comment* and locates *Yield* between *Provide Aid* (CAMEO 07) and *Investigate* (CAMEO 09). Technically, all three of these systems provide only nominal categories, and the placement of each category is arbitrary, but in fact the categories are often treated as ordinal or even interval variables. To the extent that one wishes to do that, CAMEO categories have an ordinal increase in cooperation as one goes from category 01 to 09, and an ordinal increase in conflict as one goes from 10 to 20.

Finally, we developed a formal codebook for CAMEO with descriptions of each category and examples to illustrate the types of events that fit into each category. Following the model of the IDEA codebook, the CAMEO codebook exists in both printed and web-based formats. We have also followed the lead of IDEA in introducing 4-digit tertiary coding categories that focus on very specific types of behavior, for example differentiating agreement to, or rejection of, cease-fires, peacekeeping, and conflict settlement. We anticipate that in most of our analyses the tertiary categories will not be used—we will instead aggregate the data to the secondary or primary levels—but this framework retains distinct codes for very specific behaviors that might be useful in defining patterns. It is also our expectation that these tertiary categories will assist in clarifying what event forms the respective secondary and primary categories are meant to include, thereby allowing more precise and inclusive coding.

The creation of CAMEO benefited substantially from the intellectual, experiential, ethnic, and gender diversity within the research group. Alker (1988:224)—citing a story about an Egyptian female graduate student who coded event data differently than her U.S. white male counterparts—has raised the possibility of cultural and gender biases in event data coding. In light of this concern, we note that the creators of CAMEO include a Cypriot, a Palestinian, an Iranian-American, a first generation American whose parents were born in India, and several Kansans. Some participants had extensive coding experience, others had strong knowledge of the conceptual and theoretical propositions in the field of conflict resolution, and some combined both. The regional interests of our team were similarly varied. Finally, our 2001-2002 research team had substantially more women than men (seven women and three men) and the core group responsible for most of the development was almost exclusively female. With such diversity among the makers of CAMEO, our meetings often involved intense discussions about how to make CAMEO a reasonable coding system that balanced theoretical concerns and empirical reality. Our initial disagreements usually turned to be advantageous as they directed us to do further research in an attempt to find common ground between the practitioners and the theorists.

The Mechanics of Creating New Dictionaries

When automated coding is used, the implementation of a coding system rests in the dictionaries that have been developed to associate verb phrases with the event codes in the framework. The dictionaries we have developed for CAMEO are the results of an extended process of integrating dictionaries produced by the KEDS project over the past ten years. The first step in this integration occurred before we had decided to create CAMEO. During the period 1990-1998, we had accumulated a number of different KEDS coding dictionaries, all of which used a WEIS framework. These included

Balkans and West Africa dictionaries from two independent projects (Goldstein & Pevehouse, 1997 and Huxtable, 1997 respectively), and about a dozen dictionaries that had been used to produce one-year experimental data sets on individual countries such as Mexico, China, and Russia. In addition, the Levant dictionary, which has always been the focus of KEDS project research, had been developed almost continuously during the life of the project. On the positive side, this dictionary contained the combined efforts of more than a dozen coders;[3] on the negative side, we found it retained a number of verb phrases intended solely to get around bugs and limitations in earlier versions of KEDS as well as some phrases that were added when we were still relatively inexperienced at automated coding.

Using TABARI's Merge feature, which produces a comparison of two dictionaries, we created unified 'standard' dictionaries of actors and verbs. This involved integrating the entire verb phrase vocabulary that dealt with general political behaviors and combining all general political actors (e.g., nation-states, leaders of major powers, major intergovernmental and nongovernmental organizations) from the regional actor dictionaries. We eliminated phrases involving behavior that was idiosyncratic to specific regions or crises (for example, the collapse of a pyramid investment scheme that triggered the civil disorder in Albania in 1997) as well as phrases that were either excessively long (and thus unlikely to occur more than once) or that were present in the dictionaries only because of earlier problems with KEDS. We also deleted phrases or actors that were clear when used within a single region but ambiguous in a global context.

We then used these standard dictionaries as the basis for developing new, region-specific dictionaries for the Balkans, Levant and West Africa. The first step in this process was to reincorporate regional actors into the actor dictionaries. Using the KEDS program, we went through about 18 months of *Agence France Presse* (AFP) lead sentences for each of the three regions. This new coding served the combined purposes of familiarizing new coders with KEDS, checking the standard dictionaries to make sure that we had not inadvertently deleted useful phrases, and adding vocabulary specific to AFP since all of the earlier dictionary development had been done on Reuters.

Once this process was complete, we were ready to make the transition to CAMEO. We first carried out a final manual review of the revised WEIS dictionaries to eliminate remaining problematic phrases. We then changed all of the WEIS codes that mapped directly into CAMEO using a global search-and-replace on the dictionary files. This, however, dealt with only about a third of the 4000+ verb phrases in the dictionaries. The remainder of the code changes were done manually, with pairs of coders working from printed copies of the dictionaries. Unsurprisingly, this process revealed a number of previously unnoticed ambiguities in the CAMEO codebook and we made additional changes, some quite substantial, to address these problems. Throughout this period we held weekly meetings with the entire research group to discuss potential problems with the CAMEO scheme and to identify elements of CAMEO that might prove difficult, if not impossible, to implement consistently. When the manual updating was finished, we went through the AFP leads again, this time with CAMEO rather than WEIS coding. At this point, we also made a transition from using KEDS as our coding program to using the newer TABARI program. Although we have continued to

[3] The term 'coder' refers to the individuals who are working on dictionary development in KEDS or TABARI. We use the terms 'coder' and 'coding' because they involve a third as many syllables as 'dictionary developer.' Most of our coders have been undergraduate honors students, with assistance and supervision from graduate research assistants.

make minor changes in the CAMEO codebook based on feedback from the coders, we believe that we are close to having a final version of the framework.[4]

Description of CAMEO

The main distinguishing feature of CAMEO is its explicit incorporation of event codes related to mediation and negotiation, particularly in the extended *Consult* category (CAMEO 02). For instance, *Engage in mediation* (CAMEO 025) is used when one party meets with others to play the role of a mediator. *Engage in negotiation* (CAMEO 026) is used when parties come together to negotiate, potentially to arrive at a settlement on (a) particular issue(s). Rather than assuming that all visits and meetings constitute negotiation or mediation, or inferring when events of mediation occur at the analysis stage using patterns of general visits and meetings—as we did in Schrodt et al. (2001) and Schrodt & Gerner (2001)—CAMEO allows a precise and direct distinction between mere visits and meetings on the one hand, and cases of mediation and negotiation on the other hand. Although this differentiation is clearly subject to the explicitness of the news stories, two leads show how CAMEO codes this distinction:

> Taiwan's Vice Foreign Minister *visited* Russia today, becoming the island's highest ranking government official to go there.
>
> Qatar's emir, Sheikh Hamad bin Khalifa al-Thani, *launched* a *mediation effort* on Saturday between the Emirates and Saudi Arabia whose ties have been strained by Riyadh's new friendship with Tehran.

While the first lead is coded as the linked events of *Make a visit* (CAMEO 022) and *Host a visit* (CAMEO 023), since the purpose of the visit is not clearly stated, the second lead is an unequivocal example of *Engage in mediation* (CAMEO 025) and is coded correspondingly.

Mediation-related event codes are part of the overall CAMEO framework. Subcategories that represent event types relevant to contemporary conflicts and their resolution—cease-fire, withdrawal, peacekeeping, and settlement, as well as negotiation and mediation—are present in a parallel fashion under *Request/Propose*, *Agree*, *Demand*, *Reject*, *Threaten*, and *Reduce Relations*. The following leads illustrate how such events would be coded using CAMEO:

> A group of prominent Liberians, including its foreign minister and Washington ambassador, have written to President George Bush *urging* him to *send* U.S. *peacekeeping* troops to their capital Monrovia. [Liberia *Ask for protection or peacekeeping* 054 USA]
>
> European Community foreign ministers *demanded* the *withdrawal* of Yugoslav federal forces from Bosnia-Herzegovina on Monday calling them an occupying army, diplomats quoted an EC declaration as saying. [EEC *Demand withdrawal* 106 Yugoslavia]

[4] This entire process took considerable time and effort. The original consolidation that produced the standard dictionaries took two research assistants nearly six to nine months. The development of CAMEO itself involved about six months, with between three and six people immersed in the process at various times. Conversion of the dictionaries from WEIS to CAMEO required another month and a half. (All of these times are approximate and involve student research assistants, graduate and undergraduate, working between 10 to 30 hours in a given week. Stress on everyone and public speaking demands on Gerner and Schrodt in the weeks following the attacks on 11 September 2001 undoubtedly delayed completion of CAMEO somewhat, but we had put in about two months of effort on the system even before that event.

> Palestinian leader Yasser Arafat Wednesday *rejected* a US *offer to host* a *summit* in mid-July to hammer out a framework agreement for peace between the Israelis and the Palestinians. [PLO *Reject mediation* 124 USA]

> The Soviet Union has *threatened* to *stop negotiations* to reduce long-range nuclear weapons if the United States goes ahead with the planned deployment of new medium-range nuclear missiles in Europe, the Washington post reported today. [USSR *Threaten to halt negotiations* 1311 USA]

Another significant improvement in CAMEO is our creation of four different cue categories of violence, each with various subcategories. While all forms of violence were lumped into a single and largely problematic cue category of *Force* under WEIS, four more specific and less ambiguous categories of force are created under our new framework. These categories incorporate and extend not only *Force* (WEIS 22) but also *Expel* (WEIS 20) and *Seize* (21), thereby creating conceptually coherent and codeable event forms.

CAMEO categorizes various event types that depict some form of violence into *Use structural violence* (CAMEO 17), *Use unconventional violence* (CAMEO 18), *Use conventional force* (CAMEO 19), and *Use massive unconventional force* (CAMEO 20). 'Structural violence' involves force against the rights and properties of civilians and other forms of institutionalized violence that result from the structure of socioeconomic and political relations, while 'unconventional violence' refers to physical force that does not require high levels of organization and technological sophistication. Typically, the source actors of events coded under this category are sub-state actors that are not organized and do not possess weaponry designed for sustained high levels of violence.

Use conventional force, on the other hand, encompasses uses of force and acts of war by organized armed groups, whether state or nonstate. Military blockades and occupations, and uses of weaponry ranging from small arms to artillery and to aerial bombs are coded under this cue category. Use massive unconventional force refers to the use of unconventional weaponry with massive destructive capacity, such as CBR (chemical, biological, radioactive) and nuclear weapons. The following examples, which would all have been coded under *Force* in WEIS, illustrate how CAMEO enables a more precise coding of events that demonstrate various types of force and potentially different levels of violence.

> Irish nationalist guerrillas *wounded* two British soldiers in a *bomb attack* on Thursday, police said. [IRA *Suicide, car, and other bombing* 183 Britain]

> Israel today *mounted* its long-threatened *invasion* of South Lebanon, ploughing through United Nations lines on the coast of south of Tyre and thrusting forward in at least to inland areas. [Israel *Military occupation of territory* 192 South Lebanon]

> Iraq said tonight its *warplanes attacked* Iran's main oil export terminal at Kharg Island in the Gulf and a gas plant in the southern Iranian city of Ahwaz. [Iraq *Aerial attack* 195 Iran]

In addition to correctly coding events relevant to conflict and mediation, another major theoretical objective that shaped CAMEO was to differentiate between events that have already taken place and those that may or may not occur in the future. We made sure that each subcategory present under *Request/Propose*, *Agree*, *Demand*, and *Threaten* has a corresponding code somewhere throughout the framework for use when those proposed, agreed on, demanded, or threatened actions actually take place. We wanted to assign different codes to an agreement on a certain issue on the one hand, and the implementation of that agreement on the other hand, since agreeing on or promising a particular action does not guarantee that the agreement or the promise is to be honored. This is a critical distinction, especially in mediation and conflict resolution studies, since settlements

are at least as likely to fail during implementation as they are in the negotiation phase. Expressing a commitment to negotiate or to accept mediation is similarly distinct from actually sitting at the negotiation table, and the two are coded differently under CAMEO as the following examples illustrate.

> Afghan rebel leaders *said* on Wednesday they *would meet* U.N. *mediator* Diego Cordovez if he gave them a veto over any settlement reached in peace talks. [Afghan rebels 'Agree to mediation' 06 UN]

> Arab League Secretary General Chadli Klibi, supported by Algeria and Saudi Arabia, *undertakes* a *mediation* mission between Syria and Palestinian leader Yasser Arafat. [Arab League 'Engage in mediation' 025 Syria] [Arab League 'Engage in mediation' 025 PLA]

Similarly it is important to distinguish between mere threats and the implementation of such threats. The following leads and the accompanying CAMEO codes exemplify how the new coding framework allows this distinction.

> Moscow tonight *warned* Japan it could face a *retaliatory strike if* it agreed to the deployment of more weapons aimed at the Soviet Union. [USSR 'Threaten conventional attack' 135 Japan]

> Vietnamese and Kampuchean forces *were battling* for control of a strategic base near the border today, Thai military sources said. [Vietnam 'Use of conventional force' 190 Cambodia] [Cambodia 'Use of conventional force' 190 Vietnam]

Comparison of CAMEO and WEIS Data

In this section we present a comparison of our principle data sets coded with the WEIS and CAMEO frameworks on a number of dimensions. We first compare the overall distribution of events by two-digit cue categories. We then look at the distribution of general event types over time, using monthly aggregations, to determine the extent to which the two systems pick up different behaviors. Finally, we compare a time series of mediation and negotiation events in CAMEO with a pattern-based mediation measure that we derived from WEIS data in earlier research.

Table I shows the coverage of the three data sets we have generated. The source texts were from the Reuters files on the NEXIS data service prior to 10 June 1997, Reuters Business Briefing for 11 June 1997 to 31 May 1999, and Agence France Presse (AFP) on the NEXIS data service for 1 June 1999 to 31 July 2002. The listed states in each data set correspond to the terms used in the NEXIS and Reuters searches to find the texts to be coded. Both the CAMEO and WEIS data were coded with version 0.4.04B2 of TABARI. Following the standard procedure for the KEDS project, these data sets were generated using fully automated coding, with no manual adjustments to individual records. This insures that the data generation process is completely reliable and reproducible, that the dictionaries reflect the true coding protocols, and that different coders working on various parts of the data introduce no statistical artifacts. In addition, the Levant data sets have been run through a "one-a-day" filter that eliminates multiple occurrences of a dyad-event within a single day. (In other words, if two or more events having the same source, target and event code are reported within a single day,

only one is retained.)[5] The filter reduces the size of the data sets by about 15%. While our comparisons are based on the filtered data, both the filtered and unfiltered versions are available on the web site. This analysis used the CAMEO dictionaries current on 1 May 2002. For the WEIS coding, we used the CAMEO actor dictionaries and the WEIS verb dictionaries finalized in November 2001. The latter do not incorporate additional verb phrases that we found while developing CAMEO but are otherwise relatively complete.

Distribution of Events in WEIS and CAMEO

Our first statistical comparison of the CAMEO and WEIS frameworks examines the overall distribution of events by cue categories. Tables II and III show the numerical distribution of events in the three data sets; Figures 1 and 2 show the percentage distribution. Most of the differences between the two systems conform to our expectations. For example, events in WEIS's *Force* category are distributed across CAMEO's three main violence categories, and we have substantially higher counts in the CAMEO *Agree* category because it combines several WEIS categories.

We hoped that CAMEO would reduce the number of low-frequency categories found in WEIS, but this did not occur: The standard deviation in the cue category event counts is similar in the two frameworks and in fact is slightly higher in CAMEO. Although we eliminated or combined several of the low-frequency WEIS categories (for instance, by merging *Promise* and *Agree* as well as *Warn* and *Threaten*) we introduced new low frequency categories such as *Investigate* and *Civilian Direct Action.* We included potentially low frequency behaviors—*Investigate*, for instance—because, while they are not common, the theoretical literature suggests they are important in contemporary conflict resolution. In other instances, categories have low frequencies because of the characteristics of the specific protracted disputes we are studying. This is notably in case of the *Improve Relations* category, which we introduced in order to provide symmetry with *Reduce Relations.* Relations don't improve much in these regions.

We noticed two interesting features in Figures 1 and 2 (which give percentage distributions and therefore can be compared across regions). First, the overall pattern of events is roughly similar in the three cases although they are drawn from three disparate geographical regions and, in the case of the Levant, include ten years (1979-1988) that are not coded in the Balkans and West Africa. This similarity is reassuring, given that these cases all involve protracted conflicts with substantial third party mediation and some international peacekeeping. It is particularly interesting that while the *frequency* of events in West Africa is substantially less than that in the Balkans and Levant, the overall *distribution* is similar.

Second, the pattern of events in our data—whether WEIS or CAMEO—differs substantially from that found in the ICPSR WEIS, which covers 1966-1978 and is based on *The New York Times*. The key difference is that our data include about half as many *Comment*s as the ICPSR data set and about twice as many *Consult*s (McClelland, 1983: 172). This is partly due to the characteristics of the regions we are coding. These areas have active mediation and negotiation, so we would expect to see more reported meetings than one would find in the world in general. On the other hand, we deliberately de-emphasized the use of the *Comment* category in our dictionaries for two reasons: We had difficulty differentiating neutral, pessimistic, and optimistic comments; in addition, we suspect that in many cases *Comment* may simply be an artifact of reporters or editors seeking out stories.

---

[5] We added this filter because we found the methods we used to eliminate duplicate stories in Reuters did not work well for AFP. AFP generates a large number of repeated and slightly updated stories for violent events (notably suicide bombings) as reporters obtain additional information; this was skewing the event counts. There may be occasions when this filter eliminates a second unique event. This is far less frequent than the instances of repeated stories, however, so we have chosen to err on the side of potential undercounting rather than on the side of definite overcounting.

More generally, the *Comment* category is not particularly useful. For example, the Goldstein (1992) scale for WEIS assigns comments values in the range -0.4 to +0.4 in a scale that ranges from -10.0 to +10.0, and McClelland (1983: 172-173) reports that the category was only added to WEIS as an 'after-thought.'

Comparison of Event Counts in WEIS and CAMEO

To further contrast the CAMEO and WEIS coding schemes, we compare the monthly event counts in the data sets, aggregated into four main event categories: verbal cooperation, material cooperation, verbal conflict, and material conflict. Because many of the categories in CAMEO and WEIS do not correspond directly, we use the cue category aggregations shown in Table IV. Most of our comparisons involve correlation between the monthly totals of events generated by the two coding systems; we also present selected time-series graphs of the two series. To get additional detail, we look at event totals involving several different subsets of dyads. These are shown in Table V. The "All" subset looks at the distribution of events found in any dyad in the data set. The 'conflict" subset looks at events involving the actors that are most likely to be involved in conflict with each other. Finally, the "Mediation" subset looks at events where actors that have usually been mediators are the source of action and actors involved in conflict are the targets. Our actor dictionaries code for a number of internal actors—notably ethnic groups in the Balkans and various rebel factions in West Africa—but the aggregations look only at the 3-character national code.[6]

The results of the analysis are shown in Tables VI, VII, and VIII. The comparisons across the various geographical regions are quite consistent. Virtually all of the series correlate at a very high level, usually with $r > 0.90$. Even in the case of the lower frequency aggregations involving West Africa the correlation is greater than 0.70 in all but three instances. (All of these correlations are significant at $p < 0.001$.) In two-thirds of the cases, the lowest correlations are found on events involving material cooperation: CAMEO's Provide *aid*, *Yield,* and some *Agree* categories.

Comparison of CAMEO mediation events and WEIS pattern-based mediation indicators

In a final test, we compare the events in CAMEO that specifically deal with mediation and negotiation—CAMEO event categories 025, 026, 056, 057, 058, 059, 065, 066, 068, 105, and 108—with the pattern-based measure of mediation that we derived earlier from WEIS data for the Levant and Balkans (Schrodt et al., 2001; Schrodt & Gerner, 2001). This comparison indicates how much specific mediation activity CAMEO picks up—in other words, how many of the reports mention mediation or negotiation explicitly, as distinct from just referring to generic meetings or diplomacy—and also provides a validity check on the WEIS-based measure.

In the two earlier papers, we used a simple—and somewhat indirect—indicator of mediation: the number of instances where the mediator has a cooperative interaction (WEIS categories 01 through 10) with both sides of the conflict within a period of seven days.[7] This pattern does not guarantee that the third party is actually engaged in mediation—and our future work will use more

---

[6] In our dictionaries, internal actors are coded using a three-character state code followed by a three-character code identifying the internal actor. For example, 'Liberian government' is coded LBRGOV whereas armed Liberian rebels not identified with a specific group are coded LBRREB. These identifications are particularly complicated in the Balkans, where one gets SERBS_WITHIN_BOSNIA [BFRSER], BOSNIAN_CROATS_AND_SERB [BFRSER/BFRCRO] and BOSNIA'S_WARRING_PARTIES [BFRMOS/BFRCRO/BFRSER]. Because we are using machine coding, the actor dictionaries are, in effect, the codebook for determining how various actors are identified.

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

detailed measures—but almost all mediation activities will satisfy this criterion. In other words, this measure provides a necessary but not sufficient indicator of mediation activity.

Figures 3 and 4 show the two series for the Balkans and Levant respectively; note that the CAMEO counts have been multiplied by five for the Balkans and 10 for the Levant to make the vertical scales in the figures comparable. The correlations between the WEIS and CAMEO series are $r = 0.57$ for the Balkans and $r = 0.59$ for the Levant, and both are significant at the $p < 0.001$ level. In both regions, the two series track each other quite well and there are no clear patterns with respect to the political events or crisis phases where one measure is consistently higher (or out of synchronization) with the other.

We would guess that for most applications, the measures could be used interchangeably; at the very least, their correlation reinforces the validity of each approach. The advantage of the direct CAMEO measure is that the word 'mediation' or 'negotiation' has been used explicitly in the story, insuring that the events in question are in fact events of mediation or negotiation—hence the more reserved mediation count compared to WEIS. The disadvantage of the CAMEO measure is that its identification of these phenomena depends at least in part on a potentially idiosyncratic choice of vocabulary by a reporter or editor. Explicit mediation and negotiation events are also far less frequent than the patterns of mutual meetings we looked at earlier. One reason may be that after initial stories have reported that the meetings involved 'mediation', subsequent stories simply mention that the meeting occurred (particularly when the participants refuse to say anything about what happened). The pattern-based measure would also pick up situations where an actor did not want to explicitly state that he or she was engaged in mediation, but where the sequence of consultations would be consistent with an ongoing mediation activity. Since we can still derive the pattern based measure from CAMEO—the relevant event categories remain in the framework—we will probably experiment with using both measures in the future.

Conclusion

We have summarized the reasons we developed a new coding scheme, the process we used to implement this, and some of the similarities and differences in the data sets that result. We end with some observations about how our experience relates to the more general enterprise of event data analysis.

First, it is notable that CAMEO is the third major general event data coding framework—a framework designed to categorize all types of political interactions, rather than a limited repertoire of actions such as those involving conflict—to be introduced since 1993, joining PANDA and IDEA. This follows a period of about 30 years when no new systems were introduced. Furthermore, not only have we produced a coding framework, but we also have produced data sets containing over 200,000 events coded in that framework.[8]

The difference between the earlier event data research and the current environment is, obviously, the availability of automated coding, which provides much greater flexibility and cumulativeness in dictionary development and speeds up the process of actual coding by a factor of several million times. This means that one can continue to refine a coding scheme while working on a research project. This is particularly valuable when it appears necessary to split a coding category: dictionaries can be revised relatively quickly by simply searching for the relevant code, and then determining which of the new categories each phrase should be assigned to. When human coding was

[8] The IDEA project has recently posted a massive event data set covering the entire world. This increases the amount of event data available to the research community by an order of magnitude (http://gking.harvard.edu/data.shtml).

used, it was impossible to split categories without going back through the original source texts; in practice, that was not done.

In the long run, we anticipate that event data coding schemes could evolve using a 'mix-and-match' framework whereby a researcher could adopt most of his or her coding categories from a standard set and then elaborate a smaller number of new categories. For example, a data set dealing with trade negotiation would not require any of the detail CAMEO has on cease-fires and peacekeeping and would require substantially more detail on imposition of tariffs, non-tariff barriers, and appeals to the World Trade Organization. However, primary categories such as *Consult*, *Agree*, and *Reject*, and many of the secondary categories that deal with behaviors not specific to mediation or trade would be the same. Common vocabulary of dictionaries could also be shared and the focus of the new dictionary development could be on the behaviors specific to the particular theoretical issue of interest.

Furthermore, we contend that the patterns of most political behaviors have a significant empirical component that is distinct from the theoretical considerations of the academic literature on the subject. It will be necessary, therefore, to experiment with coding systems rather than trying to establish these a priori. Due to the strong selectivity of news reports, the fact that a behavior may be important in a case study—the analytical approach that still informs most of the mediation literature—does not mean this behavior will necessarily show up as a useful statistical indicator. For example, we eliminated a number of tertiary categories in CAMEO when we were unable to find any examples for the codebook of news leads illustrating those categories. Similarly, exploratory analysis of the event data may reveal indicators not found in the theoretical literature, often because these serve as surrogates for other variables. We are not arguing that statistical studies should be atheoretical. We are saying, however, that the development of useful statistical models will, in part, be an empirical exercise of matching methods to data.

To avoid the risk of sharing McClelland's fate and being the subjects of a quote many years from now stating 'Gerner et al. never anticipated that CAMEO would still be in use in 2040...', we should make it clear that we do not consider CAMEO a definitive new event data coding framework, even for the study of third party mediation. Instead, we consider it—along with PANDA and IDEA—as an experiment in alternative ways that event data might be coded. That said, we probably have successfully done some brush-clearing in our transition away from the WEIS cue categories, notably by combining WEIS categories that could not be differentiated and eliminating categories that almost never occur in reported events.

But several problems remain. First, while we have done a lot of work on clarifying categories of verbs, we have done very little with actors. Based on our earlier experimental coding of a number of countries, we have a fairly comprehensive list of sub-state 'agents' such as police, military, judiciary, various government ministries and the like, but we have not consistently implemented these in our existing dictionaries. The coding of ethnic groups is particularly problematic—for example in some earlier Balkans data sets we put the ethnic identification of a group such as 'Bosnian Serbs' in the first three characters of the actor code, whereas the Goldstein & Pevehouse (1997) data set put the ethnic identification second (we have subsequently standardized on the Goldstein & Pevehouse convention). We have also considered the possibility of using three-part codes that would identify the nation, position, and individual (e.g. US Secretary of Defense Donald Rumsfeld would be coded as USA-DEF-RUM) with sub-fields left blank when more specific information is not available.[9] The PANDA and IDEA data sets use a separate 'agent' field rather than combining codes, and because the IDEA research group has greater experience in coding sub-state activities than we have, we will closely watch how they proceed.

[9] This would, however, require substantial additional changes in TABARI and extensive dictionary development, so we are unlikely to undertake this until we have become convinced that it is necessary for our analysis.

A second area where we still feel that our coding scheme is ambiguous involves the distinction between 'conventional' and 'unconventional' conflict. We can clearly delineate the ends of this continuum: nation-state militaries fighting World War II-style battles is conventional, and a terrorist tossing a bomb into a church is unconventional. However, there is a very large gray area—which is unfortunately becoming increasingly common in terms of behavior—involving the use of conventional military weapons against civilian populations (e.g. Israel's use of tanks and fighter aircraft to attack targets in Gaza City, Ramallah and Bethlehem) and unconventional weapons such as car bombs and suicide bombing directed against military forces (e.g. Lebanon's Hizbollah and various Palestinian militant groups in their attacks against Israel). Our inability to differentiate these activities is not confined to event data analysis—for example the USA and Israel consider Hizbollah's activities as terrorism, whereas many other states consider Hizbollah's activities legitimate opposition to a military occupation—and it is also not clear that these distinctions are theoretically important for our work on mediation. Nonetheless, having gone beyond WEIS in separating out different types of lethal activity, it is not clear how far we should go in that differentiation.

Despite these remaining ambiguities, we are optimistic that CAMEO will provide a significant improvement over WEIS. One of the first empirical studies to use the new data, Horowitz and Simpson (2002), found that the standard errors in the parameter estimate of a zero-inflated negative binomial model were substantially reduced when CAMEO was used instead of WEIS. This may or may not be repeated in all future studies but these initial results are very promising and are consistent with what we hoped to achieve with the new coding scheme.

References

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 (219-241).

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

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

Davies, John L. & Chad K. McDaniel, 1993. 'The Global Event-Data System', in Richard L. Merritt, Robert G. Muncaster & Dina A. Zinnes, eds, *International Event-Data Developments: DDIR Phase II*. Ann Arbor: University of Michigan Press (19-44).

Gerner, Deborah J. & Philip A. Schrodt, 1998. 'The Effects of Media Coverage on Crisis Assessment and Early Warning in the Middle East', in Susanne Schmeidl & Howard Adelman, eds, *Early Warning and Early Response*. New York: Columbia University Press-Columbia International Affairs Online [https://wwwc.cc.columbia.edu/sec/dlc/ciao/book/schmeidl/].

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

Gerner, Deborah J.; Philip A. Schrodt, Rajaa Abu-Jabr, Ömür Yilmaz & Erin M. Simpson, 2001. 'Determinants of "Successful" Third-Party Mediation in the Middle East: Statistical Studies of Event Data, 1979-1999', paper presented at the annual meeting of the Middle East Studies Association, San Francisco, 18 November.

Goldstein, Joshua S., 1992. 'A Conflict-Cooperation Scale for WEIS Events Data', *Journal of Conflict Resolution* 36(2): 369-385.

Goldstein, Joshua S. & 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.; Jon C. Pevehouse, Deborah J. Gerner & Shibley Telhami, 2001. 'Dynamics of Middle East Conflict and US Influence', *Journal of Conflict Resolution* 45(5): 594-620.

Horowitz, Michael and Erin M. Simpson. 2002. "Let's Make a Deal: Modeling Israeli-Palestinian Agreements, 1979-2002" Paper presented at the Summer Conference on Political Methodology, Seattle, WA, July 2002.

Howell, Llewellyn D., 1983. 'A Comparative Study of the WEIS and COPDAB Data Sets', *International Studies Quarterly* 27(2): 149-159.

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

King, Gary & William Lowe, 2001. 'An Automated Information Extraction Tool For International Conflict Data with Performance as Good as Human Coders: A Rare Events Evaluation Design', manuscript, http://gking.harvard.edu/preprints.html (accessed 19 March 2002).

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

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(2): 169-177.

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

Rasler, Karen, 2000. 'Shocks, Expectancy Revision, and the De-escalation of Protracted Conflicts: The Israeli-Palestinian Case', *Journal of Peace Research* 37(6): 699-720.

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

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

Schrodt, Philip A. & 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(3): 825-854.

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

Schrodt, Philip A. & Deborah J. Gerner, 2000. 'Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant', *American Political Science Review* 94(4): 803-818.

Schrodt, Philip A. & Deborah J. Gerner, 2001. 'Analyzing the Dynamics of International Mediation Processes', paper presented at the annual summer meeting of the Political Methodology group, Emory University, 19-21 July, and at the International Studies Association joint conference, Hong Kong, People's Republic of China, 28 July.

Schrodt, Philip A.; Deborah J. Gerner, Rajaa Abu-Jabr, Ömür Yilmaz & Erin M. Simpson, 2001. 'Analyzing the dynamics of international mediation processes in the Middle East and Balkans', paper presented at the annual meeting of the American Political Science Association, San Francisco, 1 September.

Sherwin, Ronald & Lois Van Beers, 1976. *World Event/Interaction Survey Handbook and Codebook.* Manuscript, Department of National Security Affairs, U.S. Naval Postgraduate School, Monterey, CA.

Third Point Systems, 1985. *International Political Events Database Handbook, Version 1.0*. Manuscript, Third Point Systems, Monterey, CA.

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

Tomlinson, Rodney G., 1993. World Event/Interaction Survey (WEIS) Coding Manual. Manuscript, Department of Political Science, United States Naval Academy, Annapolis, MD.

Vincent, Jack E., 1983. 'WEIS vs. COPDAB: Correspondence Problems', *International Studies Quarterly* 27(2): 160-169.

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

Table I. CAMEO Data Sets

Data Set	Time Period	Actors
Balkans	April 1989 to July 2002	Albania, Bosnia, Croatia, Kosovo, Macedonia, Montenegro, Serbia, Slovenia, Yugoslavia
Levant	April 1979 to July 2002	Egypt, Israel, Jordan, Lebanon, Palestine, Syria
West Africa	April 1989 to July 2002	Benin, Cameroon, Gambia, Ghana, Guinea, Guinea-Bissau, Ivory Coast, Liberia, Mauritania, Nigeria, Senegal, Sierra Leone, Togo

Table II. Distribution of WEIS Events by Category

Categories	Balkans	Levant	West Africa
01 Yield	1544	1790	227
02 Comment	9996	17337	2121
03 Consult	17412	39390	6109
04 Approve	2504	3849	510
05 Promise	1158	1588	235
06 Grant	2106	3616	615
07 Reward	3040	4742	1172
08 Agree	5091	6778	1401
09 Request	3834	6642	894
10 Propose	3767	5552	693
11 Reject	2480	4318	426
12 Accuse	4477	9859	977
13 Protest	745	1541	171
14 Deny	659	1530	159
15 Demand	1218	1259	161
16 Warn	1298	1617	125
17 Threaten	1295	1834	158
18 Demonstrate	1245	2519	329
19 Reduce Relations	2444	4045	762
20 Expel	384	790	221
21 Seize	2588	5097	700
22 Force	5890	15388	1036
Total	*75175*	*141081*	*19202*
Standard Deviation	*3815*	*8559*	*1272*

Table III. Distribution of CAMEO Events by Category

Categories	Balkans	Levant	West Africa
01 Comment	10337	18152	2395
02 Consult	17918	40185	6174
03 Approve	2629	4172	582
04 Improve relations	929	1587	243
05 Request	6350	10693	1331
06 Agree	7771	11652	2066
07 Provide aid	1677	2177	696
08 Yield	2619	3621	550
09 Investigate	432	1149	114
10 Demand	1531	4637	624
11 Disapprove	4873	10101	997
12 Reject	2239	4488	479
13 Threaten	1841	3075	221
14 Civilian Direct Act	451	862	122
15 Military Posture	359	960	113
16 Reduce Relations	1063	1836	297
17 Structural Violence	2528	5221	710
18 Unconventional Violence	1597	5270	354
19 Conventional Force	3937	9538	451
20 CBRN Warfare	0	0	0
Total	*71081*	*139376*	*18519*
Standard Deviation[10]	*4351*	*9175*	*1407*

[10] Not including category 20

Figure 1. Distribution of WEIS Events by Category (percent)

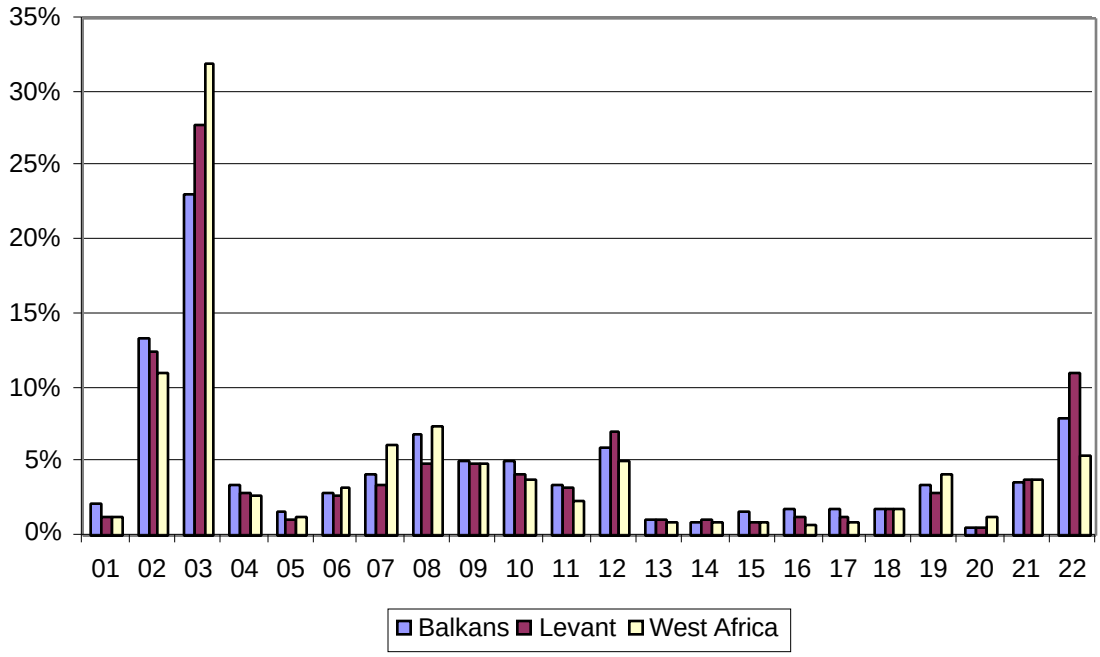


Figure 2. Distribution of CAMEO Events by Category (percent)

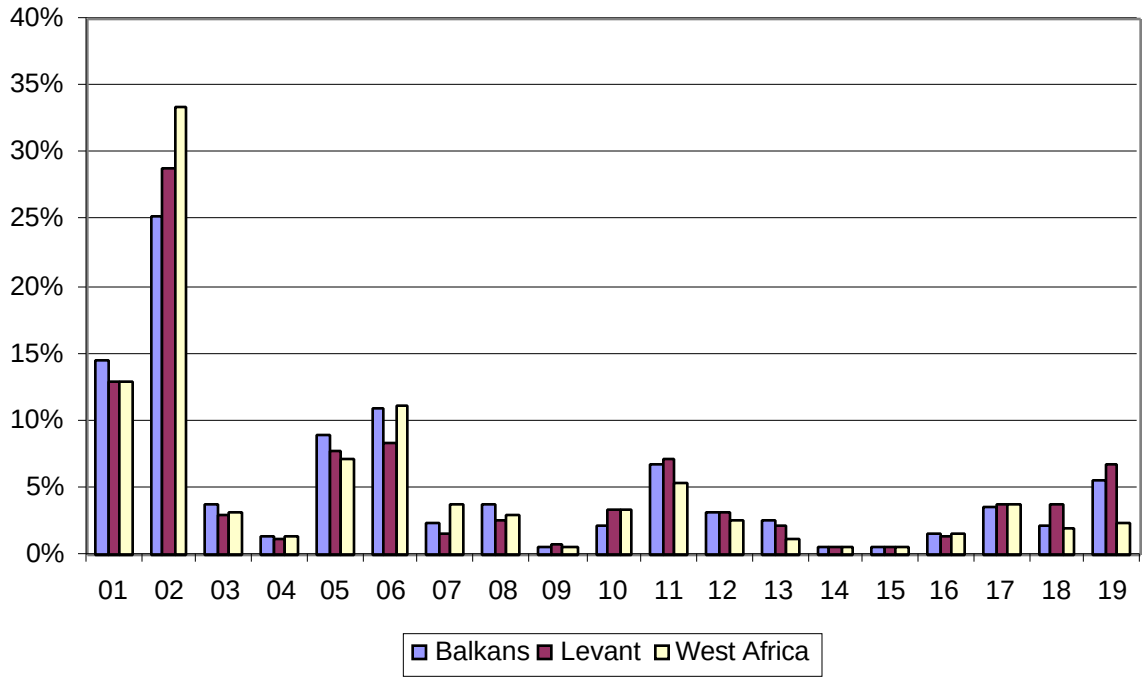


Table IV. Event Category Aggregations

Category	WEIS	CAMEO
Verbal cooperation	02, 03, 04, 05, 08, 09, 10	01, 02, 03, 04, 05, 06*
Material cooperation	01, 06, 07	063, 064, 069, 07, 08
Verbal conflict	11, 12, 13, 14, 15, 16, 17	09, 10, 11, 12, 13
Material conflict	18, 19, 20, 21, 22	14, 15, 16, 17, 18, 19, 20

* except 063, 064, 069

Table V. Dyadic Subsets

Subset	Source	Target
All	Any	Any
Conflict		
Balkans	Bosnia, Croatia, Kosovo, Serbia	Bosnia, Croatia, Kosovo, Serbia
Levant	Israel, Lebanon, Palestine	Israel, Lebanon, Palestine
West Africa	Liberia, Sierra Leone, Nigeria[11]	Liberia, Sierra Leone, Nigeria
Mediation		
Balkans	EU, France, Germany, Italy, NATO, UK, UN, USA	Bosnia, Croatia, Kosovo, Serbia
Levant	EU, France, Germany, Italy, UK, UN, USA	Israel, Lebanon, Palestine
West Africa	ECOWAS, France, OAU, UK, UN, USA	Liberia, Sierra Leone, Nigeria

[11] Nigeria was included for two reasons. First, Nigerian troops are involved in most ECOWAS military actions in Liberia and Sierra Leone, so this will pick up most of the ECOWAS intervention. Second, the data set contains quite a few reports of ethnic conflict within Nigeria.

Table VI. Comparison of WEIS and CAMEO Coding for Balkans

Events	WEIS N	CAMEO N	r
All Dyads			
Verbal Coop	43,741	45,866	0.996
Material Coop	6,714	4,353	0.963
Verbal Conf	12,181	10,920	0.986
Material Conf	12,539	9,917	0.984
Conflict Dyads			
Verbal Coop	4,610	4,728	0.969
Material Coop	968	759	0.910
Verbal Conf	1,880	1,562	0.902
Material Conf	2,954	2,457	0.953
Mediation Dyads			
Verbal Coop	6,639	6,816	0.985
Material Coop	1,381	988	0.932
Verbal Conf	2,238	2,006	0.961
Material Conf	2,271	1,637	0.953

Table VII. Comparison of WEIS and CAMEO Coding for Levant

Events	WEIS N	CAMEO N	r
All Dyads			
Verbal Coop	81,097	86,173	0.917
Material Coop	10,205	6,027	0.996
Verbal Conf	21,974	23,468	0.985
Material Conf	27,805	23,662	0.971
Conflict Dyads			
Verbal Coop	11,290	12,483	0.989
Material Coop	2,095	1,445	0.917
Verbal Conf	4,748	4,645	0.981
Material Conf	10,598	10,178	0.984
Mediation Dyads			
Verbal Coop	7,117	7,144	0.989
Material Coop	789	519	0.783
Verbal Conf	1,481	1,863	0.958
Material Conf	1,019	723	0.826

Table VIII. Comparison of WEIS and CAMEO Coding for West Africa

Events	WEIS N	CAMEO N	r
All Dyads			
Verbal Coop	11,910	12,655	0.979
Material Coop	2,051	1,318	0.692
Verbal Conf	2,171	2,445	0.913
Material Conf	3,025	2,042	0.844
Conflict Dyads			
Verbal Coop	780	823	0.908
Material Coop	124	105	0.647
Verbal Conf	213	237	0.779
Material Conf	459	279	0.902
Mediation Dyads			
Verbal Coop	672	658	0.933
Material Coop	157	136	0.661
Verbal Conf	161	182	0.751
Material Conf	235	154	0.790

Figure 3. Comparison of CAMEO mediation events and WEIS mediation patterns, Balkans

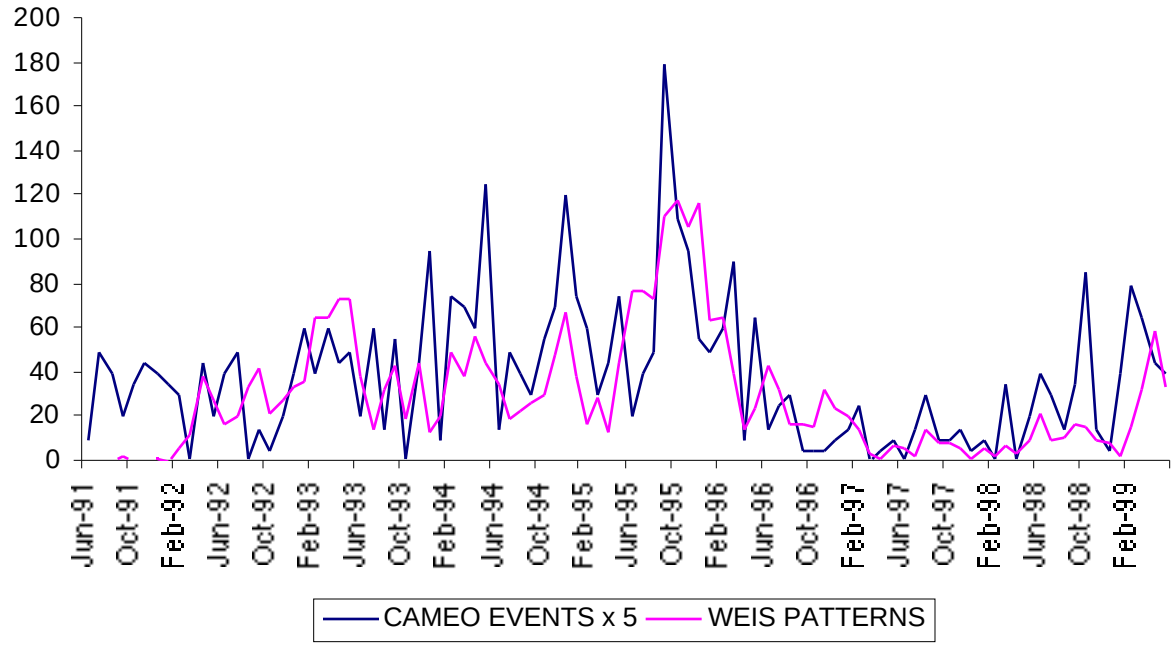
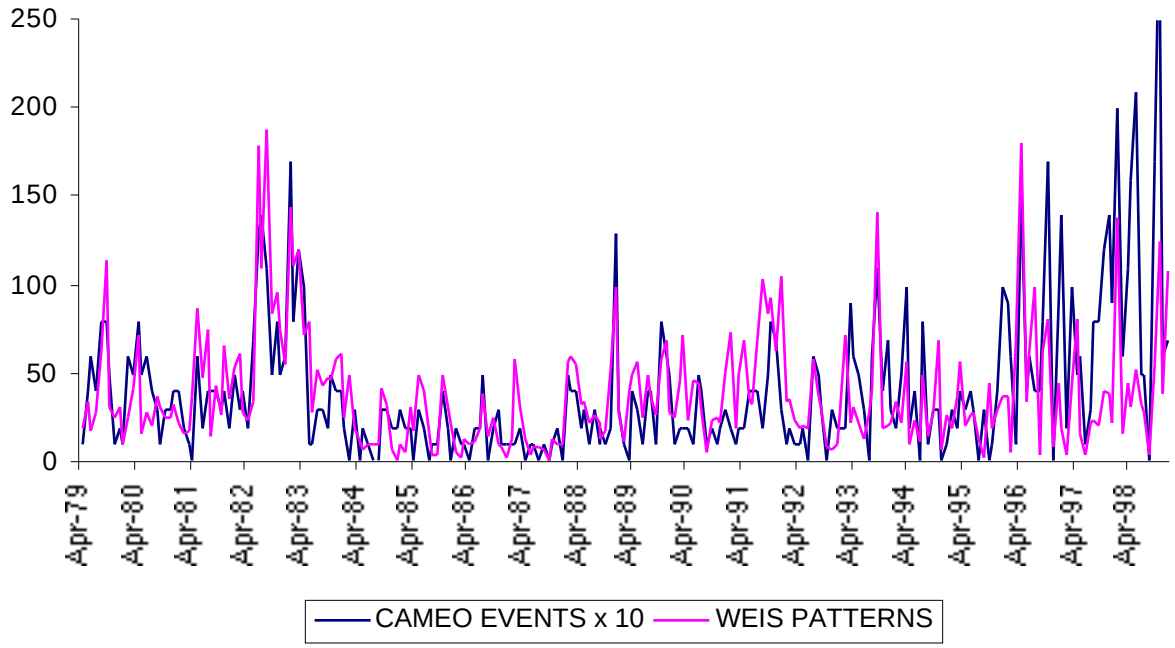


Figure 4. Comparison of CAMEO mediation events and WEIS mediation patterns, Levant



Appendix

Conflict and Mediation Event Observations (CAMEO)

01: COMMENT
- 010: Comment, not specified below
- 011: Decline comment
- 012: Make pessimistic comment
- 013: Make optimistic comment
- 014: Consider policy option
- 015: Acknowledge or claim responsibility
- 016: Make empathetic comment
- 017: Symbolic act
- 018: Announce routine activity

02: CONSULT
- 020: Consult, not specified below
- 021: Discuss by telephone
- 022: Make a visit
- 023: Host a visit
- 024: Meet in a 'third' location
- 025: Engage in mediation
- 026: Engage in negotiation

03: APPROVE
- 030: Approve, not specified below
- 031: Praise or endorse
- 032: Defend policy or action
- 033: Civilian support

O4: COOPERATE
- 040:Copperate, not specified below
- 041: Grant diplomatic recognition
- 042: Apologize
- 043: Forgive

05: REQUEST/PROPOSE
- 050: Request or propose, not specified below
- 051: Ask for information, investigation
- 052: Ask for policy support
  - 053: Ask for material aid, not specified below
  - 0531: Ask for economic aid
  - 0532: Ask for military aid
  - 0533: Ask for humanitarian aid
- 054: Ask for protection or peacekeeping
- 055: Request mediation
- 056: Request withdrawal or ceasefire
- 057: Request settlement
- 058: Request to meet or negotiate
- 059: Propose to mediate

06: AGREE
- 060: Agree, not specified below
- 061: Sign formal agreement
- 062: Agree to policy support
- 063: Agree to provide material support, not specified below
  - 0631: Agree to provide economic support
  - 0632: Agree to provide military support
  - 0633: Agree to provide humanitarian support
- 064: Agree to peacekeeping
- 065: Agree to mediation
- 066: Agree to mediate
- 067: Agree to yield
- 068: Agree to meet or negotiate
- 069: Agree to settlement

07: PROVIDE AID
- 070: Provide aid, not specified below
- 071: Provide economic aid
- 072: Provide military aid
- 073: Provide humanitarian aid
- 074:Grant asylum

08: YIELD
- 080: Yield, not specified below
- 081: Ease non-force sanctions, not specified below
  - 0811: Ease administrative sanctions
  - 0812: Ease economic boycott or sanctions
  - 0813: Ease civilian boycott or strike
- 082: Ease, stop military blockade
- 083: Return, release, not specified below
  - 0831: Return, release person(s)
  - 0832: Return, release property
- 084: Ceasefire, observe truce
- 085: Demobilize armed forces
- 086: Military retreat or surrender

09: INVESTIGATE
- 090: Investigate, not specified below
- 091: Investigate crime, corruption
- 092: Investigate human rights abuses
- 093: Investigate military action or war crimes

10: DEMAND
- 100: Demand, not specified below
- 101: Demand information, investigation
- 102: Demand policy support
- 103: Demand aid
- 104: Demand protection, peacekeeping
- 105: Demand mediation
- 106: Demand withdrawal
- 107: Demand ceasefire
- 108: Demand meeting, negotiation
- 109: Demand rights

11: DISAPPROVE
- 110: Disapprove, not specified below
- 111: Criticize or denounce
- 112: Accuse
- 113: Official protest

12: REJECT
- 120: Reject, not specified below
- 121: Reject proposal, not specified below
  - 1211: Reject ceasefire
  - 1212: Reject peacekeeping
  - 1213: Reject settlement
- 122: Reject request for material aid
- 123: Reject proposal to meet, discuss, negotiate
- 124: Reject mediation
- 125: Defy norms, law
- 126: Reject accusation, deny responsibility
- 127: Veto

13: THREATEN
- 130: Threaten, not specified below
- 131: Threaten non-force, not specified below
  - 1311: Threaten to halt negotiations
  - 1312: Threaten to halt mediation
  - 1313: Threaten to reduce or stop aid
  - 1314: Threaten to boycott or embargo
  - 1315: Threaten to reduce or break relations
- 132: Give ultimatum
- 133: Threaten blockade
- 134: Threaten occupation
- 135: Threaten conventional attack
- 136: Threaten unconventional attack
- 137: Threaten massive unconventional attack

14: CIVILIAN DIRECT ACT
- 140: Civilian direct action, not specified below
- 141: Demonstration
- 142: Hunger strike
- 143: Strike/boycott
- 144: Physical obstruction
- 145: Violent protest, riot

15: MILITARY POSTURE
- 150: Military posturing, not specified below
- 151: Military demonstration, display
- 152: Military alert
- 153: Military mobilization

16: REDUCE RELATIONS
- 160: Reduce relations, not specified below
- 161: Reduce or break diplomatic relations
- 162: Reduce or stop aid, not specified below
  - 1621: Reduce or stop economic assistance
  - 1622: Reduce or stop humanitarian assistance
  - 1623: Reduce or stop military assistance
  - 1624: Reduce or stop peacekeeping
- 163: Halt negotiations
- 164: Halt mediation
- 165: Impose embargo, boycott

17: USE STRUCTURAL VIOLENCE
- 170: Use of structural violence, not specified below
- 171: Violence against property, not specified below
  - 1711: Confiscate property
  - 1712: Destroy property
- 172: Administrative sanctions, not specified below
  - 1721: Impose curfew
  - 1722: Impose censorship
- 173: Arrest and detention
- 174: Expel, not specified below
  - 1741: Expel diplomat(s)
  - 1742: Expel group(s)

18: USE UNCONVENTIONAL VIOLENCE
- 180: Use of unconventional violence, not specified below
- 181: Abduct, hijack
- 182: Non-lethal physical assault, not specified below
  - 1821: Sexual assault
  - 1822: Torture
- 183: Suicide, car, and other bombing
- 184: Murder or political assassination

19: USE CONVENTIONAL FORCE
- 190: Use of conventional force, not specified below
- 191: Military closure or blockade
- 192: Military occupation of territory
- 193: Small arms and light weapons attack
- 194: Artillery and tank attack
- 195: Aerial attack

20: USE MASSIVE UNCONVENTIONAL FORCE
- 200: Massive unconventional force, not specified below
- 201: CBR attack
- 202: Nuclear attack

World Event Interaction Survey (WEIS)

Cue code	*Secondary code*		*Goldstein scale value*
01	**YIELD**		
	011	Surrender, yield to order, submit to arrest, etc.	0.6
	012	Yield position; arrest; evacuate; involves actual physical movement	0.6
	013	Admit wrongdoing; retract statement	2.0
02	**COMMENT**		
	021	Explicit decline to comment	-0.1
	022	Comment on situation-pessimistic	-0.4
	023	Comment on situation-neutral	-0.2
	024	Comment on situation-optimistic	0.4
	025	Explain policy or future position	0.0
03	**CONSULT**		
	031	Meet with at neutral site; or send note	1.0
	032	Visit; go to	1.9
	033	Receive visit; host	2.8
04	**APPROVE**		
	041	Praise, hail, applaud, condolences, ceremonial saluations	3.4
	042	Endorse other's policy or position; give verbal support	3.6
05	**PROMISE**		
	051	Promise own policy support	4.5
	052	Promise material support; human or resourcer aid forthcoming	5.2
	053	Promise other future support action	4.5
	054	Assure; reassure; expressions/reiterations of promise of earlier pledges	2.8
06	**GRANT**		
	061	Express regret; apologize	1.8
	062	Give state invitation	2.5
	063	Grant asylum; annoucement of a policy and reports of granting of refuge	-1.1
	064	Grant privilege, diplomatic recognition; etc	5.4
	065	Suspend negative sanctions; truce	2.9
	066	Release and/or return persons or property	1.9
07	**REWARD**		
	071	Extend economic aid (as gift and/or loan)	7.4
	072	Extend military assistance; men, material, joint military training exercises	8.3
	073	Give other assistance	6.5
08	**AGREE**		
	081	Make substantive agreement	6.5
	082	Agree to future action or procedure; agree to meet, to negotiate	3.0
09	**REQUEST**		
	091	Ask for information	0.1
	092	Ask for policy assistance	3.4
	093	Ask for material assistance	3.4
	094	Request action; call for	-0.1
	095	Entreat; plead; appeal to; help me; requests from a distinctly suppliant position	1.2

10	**PROPOSE**		
	101	Offer proposal	1.5
	102	Urge or suggest action or policy	-0.1
11	**REJECT**		
	111	Turn down proposal; reject protest, threat, etc.	-4.0
	112	Refuse; oppose; refuse to allow	-4.0
12	**ACCUSE**		
	121	Charge; criticize; blame; disapprove	-2.2
	122	Denounce; denigrate; abuse	-3.4
13	**PROTEST**		
	131	Make complaint (not formal)	-1.9
	132	Make formal complaint or protest	-2.4
14	**DENY**		
	141	Deny an accusation	-0.9
	142	Deny an attributed policy, action, or position	-1.1
15	**DEMAND**		
	151	Issue order or command, insist; demand compliance, etc.	-4.0
16	**WARN**		
	161	Give warning	-3.0
17	**THREATEN**		
	171	Threat without specific negative sanctions	-4.4
	172	Threat with specific nonmilitary sanctions	-5.8
	173	Threat with force specified	-7.0
	174	Ultimatum; threat with negative sanctions and time limit specified	-6.9
18	**DEMONSTRATE**		
	181	Nonmilitary demonstration; to walk-out on; marching, picketing, stoning, etc.	-5.2
	182	Armed force mobilization, exercise and/or displays not included here	-7.6
19	**REDUCE RELATIONSHIP (as negative sanctions)**		
	191	Cancel or postpone planned event	-2.2
	192	Reduce routine international activity; recall officials; embargos, bans, etc.	-4.1
	193	Reduce or cut off aid or assistance	-5.6
	194	Halt negotiations	-3.8
	195	Break diplomatic relations	-7.0
20	**EXPEL**		
	201	Order personnel out of country	-5.0
	202	Expel organization or group	-4.9
21	**SEIZE**		
	211	Seize position or possessions; also military occupation	-9.2
	212	Detain or arrest person(s)	-4.4
22	**FORCE**		
	221	Non-injury destructive act, including demonstrations with physical destruction	-8.3
	222	Nonmilitary injury; destruction; terrorist bombings	-8.7
	223	Military engagement	-10.0

Source: McClelland 1976; Goldstein 1993