THE MACHINE-ASSISTED CREATION OF HISTORICAL EVENT DATA SETS: A PRACTICAL GUIDE

G. DALE THOMAS

Department of Government and International Studies
Gambrell Hall
University of South Carolina
Columbia, SC 29208
(803) 777-2884 or 777-3109
THOMAS628@AOL.COM

ABSTRACT

As the title of this project suggests, the purpose of this paper is to present a practical guide to the machine assisted creation of historical event data sets. In an effort to make this useful to as large an audience as possible, I begin by introducing event data as commonly understood in international relations research and discuss why some scholars are arguing that this type of data is presently becoming more attractive. After a somewhat brief discussion of the foregoing, I turn to a series of practical steps specifically oriented toward the machine assisted creation of historical event data sets. Throughout this discussion, I will be relying on examples drawn from my experience in creating the Northern Ireland Systemic (NIS) data set which focuses on domestic political violence from 1968 to 1996. Finally, I present a brief description of work being carried out using the NIS data.

This paper was prepared for the International Studies Association Annual Meeting, March 14-18, 2000 in Los Angeles, California. The author would like to thank the Department of Government and International Studies and the Graduate School of the University of South Carolina for dissertation funding used to create the Northern Ireland Systemic Data Set.

As the title of this project suggests, the purpose of this paper is to present a practical guide to the machine assisted creation of historical event data sets. In an effort to make this useful to as large an audience as possible, I begin by introducing event data as commonly understood in international relations research and discuss why some scholars are arguing that this type of data is presently becoming more attractive. After a somewhat brief discussion of the foregoing, I turn to a series of practical steps specifically oriented toward the machine assisted creation of *historical* event data sets. I emphasize the time component of this project for two major reasons. First, a variety of articles and books have been published in the last few years regarding the creation of event data sets oriented toward the present or recent past (e.g., Schrodt and Gerner 1994; Gerner, Schrodt, Francisco, and Weddle 1994; Merritt, Muncaster, and Zinnes 1993; Bond, Jenkins, Taylor, and Schock 1997; Duffy 1994; Davies and McDaniel 1994; Mallery 1994; Woolley 2000); thus, the subject has been well covered. And second, trying to create a historical event data set involves certain considerations either not present or of less importance in typical event data set creation. Throughout this discussion, I will be relying on examples drawn from my experience in creating the Northern Ireland Systemic (NIS) data set which focuses on domestic political violence from 1968 to 1996.[1] Finally, I present a brief description of work being carried out using the NIS data.

EVENT DATA

In some ways, asking what is event data is like stepping into a definitional and philosophical quagmire. Enough ground seems to be shared by interested scholars to suggest that event data results from the nominal or ordinal coding of events. Ah, but what then is an event? Here one begins to get on shakier ground, but, loosely speaking, the actions and reactions of political actors have been termed events. However, these activities must be "reported in some reputable and available public source" (Azar 1980, 146) in order to qualify. As some authors have argued, "Events . . . can be regarded as (1) meaningful signals or communications between nations . . . (2) part of a complex stream of historical sequences . . . or (3) acts of pressure, attempts at exercising influence, or acts of yielding to it . . ." (Azar, Bennett, and Sloan 1974, 36). The two best known event data projects are best characterized by (1) and (2) of the above. Azar's Conflict and Peace Data Bank (COPDAB) follows the first, while McClelland's World Event Interaction Survey (WEIS) is based on the second.

More formally one can define an event as:

> . . . an interaction, associated with a specific point in time, that can be described in a natural language sentence that has as its subject and object an element of a set of *actors* and as its verb an element of a set of *actions,* the contents of which are transitive verbs (Gerner, Schrodt, Francisco, and Weddle 1994, 95 [emphasis in original]).

---

1 For practical reasons that will be discussed, the data set excludes 1991 and part of 1992. Furthermore, actors in the Northern Ireland conflict are under identified. Reports of many of the more violent events in the conflict do not include a clearly identified actor. Rather than excluding a large part of the interaction occurring in the conflict, I have chosen to look at the conflict as a whole rather than as a series of dyads.

This definition stresses the importance of natural language, actors, actions, and time. Since event coders do not witness the action themselves, they must code the event's reporting, which is in the form of a natural language such as English. "Empirically, events can only be defined with respect to a human language or set of languages. The event coding exercise converts the natural language into nominal data that can be analyzed using formal methods" (Gerner, Schrodt, Francisco, and Weddle 1994, 95).[2] For most event data research, the actors are political entities, but this is not required by the definition. "Any model of political activity will be specific to certain persons, organizations, and places, all of which are specified by noun phrases in the language or languages used in the source"(Gerner, Schrodt, Francisco, and Weddle 1994, 95).

Although event data records the actions and reactions of different actors, only certain types of behavior are actually represented in the coding scheme--e.g., the WEIS data is coded into 63 nominal categories, while COPDAB uses 16 ordinal categories on a conflict-cooperation scale (Schrodt and Gerner 1994, 828). Gerner, Schrodt, Francisco, and Weddle explain:

> All of these interactions can be described by transitive verbs; for example, *apologize, met with, endorsed, promise, accuse, threaten,* or *attack.* As with the nouns, multiple verbs might signify the same category of behavior, either because the words are synonyms within the language (e.g., *grant, bestow, contribute, donate, fund, present, provide*) or because the behaviors, although linguistically distinct, are politically equivalent, a characteristic that Most and Starr (1989: chap. 5) refer to as "foreign policy substitutability." These equivalence sets will vary with the specific problem or the theoretical approach and in large part determine the validity of a particular coding scheme (1994, 96 [emphasis in original]).

Finally, all event data must include time of the event. The actual unit of time depends upon the research question. Although most event data are recorded on a daily basis, the events are often aggregated into weekly, monthly, and even yearly totals for further analysis.

This brief discussion of event data gives the previously uninformed reader an idea of what event data is, but why would one want such a monster. In 1974, event data was one of the most commonly used sources in the quantitative study of international relations (Azar, Bennett, and Sloan 1974). One reason for event data's popularity was that such data provides "a convenient window onto discrete actions and communications directed from one actor to another over time" (Gerner, Schrodt, Francisco, and Weddle 1994, 92), and perhaps more importantly, event data acts as an important link between journalistic text and analysis for events occurring over long periods (Gerner, Schrodt, Francisco, and Weddle 1994). Yet, if event data can play such an important role in international relations and political science why doesn't it receive the same stress in graduate programs and professional journals that it received twenty-five years ago?

At least two major reasons exist for the decline in prestige and importance of event data in political science and international relations. The first reason is the failure of many of the early

---

2 Schrodt (1994, 38) defines an event coding scheme as a "function E that maps text strings into codes [E: *S* _ Pow(*C*)] where Pow(*C*) is the power set of *C* (all possible subsets of *C*)."

event data projects to produce expected results. Schrodt (1997) argues that event data research went into a severe decline due to the poor performance of linear early-warning systems based on event data. Simply put, the high expectations of the early 1970's were not met with significant enough success to justify the expenditures necessary to support the large event data programs. The second reason can be found in the strongly critical assessments of event data validity. These critiques frequently emerged from within the event data scholar community (e.g., Burgess and Lawton 1972; Vincent 1983).[3] One of the most frequent and telling critiques of event data is that severe problems exist in asserting that event data represents or captures an underlying objective reality (see Duffy 1994; Davies and McDaniel 1994).[4] In response to internal (and external criticism) and the failure of major research programs to achieve expected breakthroughs, many researchers began to turn to alternative forms of data collection and representation.

Nonetheless, as a number of researchers will attest (e.g., authors found in the Merritt, Muncaster, and Zinnes 1993 edited volume), event data research certainly hasn't died. Actually, quite the opposite appears to be occurring at the present. Phenomenal advances in computing power have enabled a renaissance in event data research. On one hand, the machine coding or assisted coding of event data has allowed for the relatively low cost updating of older data sets and the creation of new data sets oriented toward specific problems (e.g., see Merritt, Muncaster, and Zinnes 1993; Davies and McDaniel 1994; Bond, Jenkins, Taylor, and Schock 1997; and Schrodt and Gerner 1994). On the other hand, additional computing power allows for new types of analyses to be conducted that are not based on linear models (Schrodt 1994; 1997a; 1997b; Duffy 1994; Schrodt and Gerner 1997). Much of the rest of this project focuses on how these gains in computing power can be used to create historical event data sets at a relatively low costs both in terms of money and in terms of work-hours.

PRACTICAL STEPS TO CREATING MACHINE-ASSISTED HISTORICAL EVENT DATA

This section examines nine major steps in creating a historical event data set. Some of these steps may at first seem trivial; however, after two years of trial and error with more errors and trials than I care to remember, I suggest that each step deserves significant attention in the order listed below.

1. *Choosing a research question.* Many may feel that this step is an unnecessary addition to this discussion; however, it is not. By no means do I suggest that people should do research like a child with a new hammer looking for anything to hit it with. Event data is inappropriate for most research questions. Nonetheless, a few research questions can only be answered with this type of data. A variety of outstanding work has been written on helping one formulate an appropriate question,[5] and I do not seek to duplicate this effort here. Yet, I do believe that one should be aware of what types of questions can be more reasonably answered with event data.

---

3 For more recent critiques see Schrodt (1994), Duffy (1994), Alker (1993), and Woolley (2000).

4 This remains an important issue in the creation of historical event data sets but can be handled in a defensible and relatively simple manner during the development of the research question. This issue will be discussed at some length in the practical steps section.

5 For example, see Lave and March (1975); Gurr (1972); and Most and Starr (1989).

Of the validity issues facing event data research, one of the most important is the question of "whether any record or limited combination of records (news sources, archives, etc.) can be accepted of valid, objective representations of "real world" events . . ." (Davies and McDaniel 1994, 56). One way of dealing with this challenge is to ask what then do event data represent. This has been a major research program for a number of authors (Alker 1993; Duffy 1994; Davies and McDaniel 1994). Nevertheless, those who challenge event data on this basis have not reached a consensus on an appropriate response. Davies and McDaniel (1994) argue that one should not look at event data as representing a sample or totality of the "real" events one is studying, rather one should make the events themselves the object of study. Their theory is "built on the crucial observation that language use is not so much a passive exercise in representing the 'real world,' as a means of actively participating in the ongoing creation of overlapping social realities. . ." (Davies and McDaniel 1994, 57). Duffy (1994) focuses on the differences in how actions are perceived and create different world versions. Thus, the question of what event data actually represents remains "over determined."

I believe that whether or not one accepts the "naive realist assumption that 'real' world events can and should be identified, studied and understood (at least largely) independently of their perception and representation in news and other reports" (Davies and McDaniel 1994), one's research question affects and is affected by one's position on this issue. One may find that one cannot answer a question as originally conceived without accepting certain epistemological and ontological baggage associated with a particular view of event data. I will come back to this issue in my discussion of successfully choosing a source or sources.

In the case of the NIS data, I developed the data set in an attempt to answer the question of whether or not crisis can be identified solely on the basis of flow of interactions among conflict participants in Northern Ireland. Does crisis have some "objective foundation" or is crisis merely a subjectively identified phenomenon based on the opinions of key actors and observers? Even if the NIS data cannot be considered representative of "real" events independent of those writing the news articles or reports, I can argue that it is representative of the understanding of actions in Northern Ireland as conditioned by socio-cognitive reality of a large section of the Northern Ireland population. I believe that this is sufficient for answering the above research question. Yet, once one has chosen a research question for which event data seems appropriate, how should one proceed?

2. *Choosing a time and geographical area.* Although this does not involve a great deal beyond that associated with choosing an appropriate research question, one must still make this determination before moving on to the much more challenging issue of choosing an appropriate source. How wide does the geographic coverage need to be in order to answer the proposed research question? How much time needs to be covered? These issues will likely limit available sources. In the case of the NIS data set, the conflict in Northern Ireland began in 1968 and continued in "hot" form until 1996. Thus, any source that I might choose needed to give good coverage for as much of this time period as possible; furthermore, the conflict involved not only Northern Ireland, a region of the United Kingdom, but also the U.K. as a whole and the Republic of Ireland. Other geographical actors were also involved, however, on a much more limited basis. Therefore, whatever source that I would choose would also have to give significant coverage to how the conflict was affecting each of the major actors, not just a single actor.

3. *Finding a source.* Choosing an appropriate source remains one of the most important and troublesome issues in event data research (Azar 1980; Davies and McDaniel; Howell 1983; Huxtable and Pevehouse 1996; Vincent 1983;Woolley 2000). Some of the most popular sources for event data research have been the *New York Times* (McClelland 1976; Lebovic 1994) and the *Reuters* news service (Schrodt and Gerner 1994; Schrodt and Gerner 1997; Davies and McDaniel 1994). An alternative approach has been to use a variety of sources; COPDAB used a combination of over 70 sources (Azar 1980). Nonetheless, one must return to the important issue of what does event data actually represent when choosing a source.

Even a nuanced view of the "realist" position leads a researcher into difficulty. Two of the largest event data projects, WEIS and COPDAB, both share a certain outlook on the definition of event. McClelland and Hoggard (1969) make the distinction between transactions and interactions, with transactions being the routine exchanges between states, and interactions being the non routine exchanges between states that are often handled at the highest levels of government. These non routine exchanges are expected to be considered newsworthy enough to be reported by a reliable media source. Azar also uses the concept of "transactions" and similarly excludes the routine from his concept of event:

> These events are occurrences between nations which are distinct enough from the constant flow of "transactions" (trade, mail flow, travel, and so on) to stand out against this background as "reportable" or "newsworthy." Thus, to qualify as an "event," an occurrence has to be actually reported in some reputable and available public source. For example, the conclusion of a trade agreement would qualify as an "event," but the subsequent individual and *routine* trade exchanges conducted under its terms would not (Azar 1980, 146).

Consequently, if one considers the frequency of the appearance of a particular action in a reputable news source, one can reasonably posit that it is a function of the "newsworthiness" of that class of actions and the number of times that that action occurred during the reporting period:

$$\text{Number of Articles}=F(\text{newsworthiness of action class, frequency of action during reporting period}) \tag{1}$$

Newsworthiness can be considered a function of the frequency of an event. The more that a particular kind of event takes place, the more people are desensitized to its occurrence and therefore find it less interesting. If one defines the newsworthiness of an action class as the probability that an event in that action class will be reported if it occurs, one ends up with the following equation:

$$N_{\text{Class, Articles}} = P_{\text{Class}} * N_{\text{Class, Actions}} \tag{2}$$

For illustrative purposes, one can look at the number of articles that result from a series of actions over an arbitrary period of time by setting the maximum frequency of an action to 100 and the probability that a class of action will be reported to ($1-N_{\text{Class, Actions}}/100$). This results in the following:

$$N_{Class,\ Articles} = (1-N_{Class,\ Actions}/100) * N_{Class,\ Actions} \quad (3)$$

If these values are calculated for each possible frequency ranging from 0 to 100, the resulting values show a concave curvilinear relationship. See Figure 1.

<< Figure 1 About Here>>

This curve demonstrates an important facet of event data analysis: without knowledge of the "real" number of events one cannot know where on the curve a particular class of actions can be found, neither can one accurately gauge the population of a particular action class by reading a newspaper without knowing the probability of a particular action class being reported. Consequently, if the class of action is presidential meetings with advisors, simply by reading the newspaper one cannot know the number of times that the president met with advisors. Alternatively, if the class of action is military clash, without the outside information that two countries are at war or peace, one cannot know the frequency that two states' soldiers are shooting at one another simply by reading a news article that only covers a single military clash. These comments have totally ignored the normal questions regarding event data source bias which are valid concerns in assessing data set validity.

My belief is that many of the scholars working on early event data projects thought that they had escaped this issue by focusing on a particular level of analysis, the state. The dominance during much of the event data analysis heyday of the realist paradigm with its state-centrist outlook and focus on military issues has undoubtedly contributed to this point of view: "high politics" are always newsworthy, although they differ in the probability that they will be reported based on the geographical area in which they occur[6]. Yet, I would argue that "high politics" are not always newsworthy in the same sense. High level NATO meetings do not receive the same attention in the news media that they would if the same government personnel were to meet in the absence of a grand strategic alliance. Indeed, even "high politics" may or may not be eventful in the sense of Azar's definition of event: an event can be defined as the actions and reactions of political actors reported in a reputable and public source (Azar 1980). Thus, the relationship between actions/reactions and events is dynamic rather than static. The population of newsworthy events changes through time in direct relation to the relative frequency of actions in the relatively near past. Newsworthiness acts as a filter through which actions and reactions pass or fail to pass on their way to "print". Editor and geographical biases further filter actions and reactions leading to a representation of the "event population" that may be substantially removed from the "action/reaction" population.

Are there any solutions to this difficulty? I would say yes, although I think that the preceding line of reasoning can substantially undermine the recognized validity of both the

---

6 This assumption remains common in event data research and must be addressed if research results are to be given sound footing. Woolley (2000) reports the results of a number of studies testing this assumption. Duffy (1994) addresses this issue obliquely, but neither Woolley nor Duffy bring the test results and their justification to a satisfactory synthesis. See also Huxtable and Pevehouse (1996).

COPDAB and WEIS data sets for most dyads.[7] I call this problem Azar's paradox since Azar frequently sought to identify normal relations for a dyad while supposedly ignoring the most normal behavior for the dyad.[8] If one is looking at a neutral dyad and if the data is scaled on a conflict–cooperation continuum with neutrality being given a zero scale score, Azar's method may be successfully employed. The reason for this is that the most commonly occurring behavior, that of neutrality, which has trouble making it through the newsworthiness filter to appear in a reputable and public source, would only be zeroed out in the process of scaling and aggregating. Consequently, Azar's method, *ceteris paribus*, works if one scales and aggregates data in studying dyads that are essentially neutral.

This leaves the question of what to do when dyads are not neutral but tend rather toward the cooperative or conflictual ends of the continuum, or when one does not wish to scale and aggregate the data. My own answer, which I feel does not violate the core of the event data research program, involves using a "native" source.

Beyond the most fundamental levels, reality is a social construction through which individuals act on the basis of their subjective understanding of their circumstances. To illustrate this, consider "war". "War" is part of the socially-constructed reality of most societies. "War" can be objectively defined, as done by Singer and Small (1972), but the concept also has such strong roots in the social-construction of Western civilization that newspaper writers and lay people alike have little difficulty in identifying one if they see it in progress. "War" is also a collective activity, and although the phrase "a one man war" has achieved popularity, it is nonetheless a metaphorical usage of "war". The word "collective" presupposes the existence of the individual. "War" is collective not only in its actors but also in its actions: one fired gun does not a war make. In summary, "war" in the abstract that is socially-constructed-reality has meaning and acts as a cognitive short-hand for its individual actors and actions.

Nevertheless, the cognitive reality of which cognitive shorthands are a part may be substantially different from one area to another. If one were to ask about the implications of the former "war" in Bosnia, one would likely receive significantly different responses from a Bosnian widow who lost her husband and children in the fighting and a U.S. State Department analyst. *Thus, for event data to be "valid" it must tap into the cognitive reality of the dyadic actors.*

---

7 Criticizing COPDAB and WEIS is in some ways a wasted effort. The two data sets have been so heavily criticized that one more validity challenge makes little difference. Researchers who have questions that can be readily answered with WEIS or COPDAB data will almost certainly continue to use it and understandably so. Use of questionable data is better than not using any data at all.

8 For a Azar's introduction to a normal relations range, see Azar (1972). The fundamental idea is to scale and aggregate the event data for a dyad; then one takes the mean of the values and places a critical threshold one standard deviation above and one standard deviation below the mean. Any values falling between the two critical thresholds are considered normal behavior for that dyad.

Just as many items suffer from declining utility as their numbers increase, the cognitive effect of certain actions also declines. If one compares the pairs of a day on which a single bomb exploded with a day on which no bombs exploded and a day on which 50 bombs exploded as opposed to a day on which 49 bombs exploded, the physical reality could be the same for that single bomb and the 50th bomb, but the importance of the 50th bomb to the cognitive reality of those not directly affected is likely to be much less than it is for the single bomb. Consequently, within the cognitive reality of a society, if the probability of a class of events being reported by a “native” source is not one, a researcher can, *ceteris paribus*, reasonably assume that those actions/reactions that are of interest to the normal events data project and do not make it into a reputable source have been cognitively subsumed into the broader social reality or simply been excluded due to the more esoteric problems associated with event data.

In summary, events are signposts to the cognitive and social realities of actors who are the producers and consumers of their respective sources. Metaphorically, a large number of actions/reactions pass through a three level sieve: cognitive-social-reality, newsworthiness (which is, itself, largely dependent upon cognitive-social-reality), and editor and coverage biases. Indeed, one might perceive the process as a funnel where many actions/reactions go in but only a few events come out. The *New York Times* may be an outstanding source for event data if your interest is in U.S. national politics or directed U.S. actions but is of little use in understanding the cognitive-social-reality of Northern Ireland.

On this basis, the NIS data set has been constructed using three “native” chronologies for Northern Ireland. The three sources that I have used are *Fortnight* magazine, a *Fortnight* magazine edited chronology, and a highly acclaimed chronology by Deutsch and Magowan (1973). *Fortnight* magazine, published every two weeks in Belfast, includes a summary of events for the preceding two weeks containing political, economic, and cultural happenings deemed relevant to the magazines reading audience. The magazine is considered a reputable source and is therefore appropriate for identifying “events.” The original source material, although highly informative, contains many events that are not relevant for the question at hand and therefore requires extensive editing to isolate the primarily political events occurring during the conflict. *Fortnight* has also compiled an edited chronology (Bell, Johnstone, and Wilson 1992) that filters out many of the non political events. Given that time and money limit research opportunities, I have chosen to use the already edited chronology as the primary source for the entire period that it covers, September 1970 through December 1990. I have then supplemented this with the highly acclaimed chronology compiled by Deutsch and Magowen for January 1968 through August 1970 and by the unedited Fortnight chronology from November 1992 through May 1996. Therefore, events for the “Troubles” in Northern Ireland are covered from January 1968 through December 1990 and then November 1992 through May 1996.

Although I would prefer to have an uninterrupted chronology, two practical reasons have kept me from accomplishing this. First, the Fortnight chronology from November 1992 through May 1996 has been available on the Internet in electronic format. This is an *important* consideration when attempting to machine code event data. If one only has a paper source, several additional steps will be required in order to get the information into an electronic format. Second, none of the fifteen key-events that my study examines as crisis situations took place in the intervening months.

If at all possible, one should attempt to find a source that is available in electronic format. However, when doing historical research, this is not a high probability option. Converting a paper source to an electronic source is a fairly straightforward yet time consuming process. Optical character recognition (OCR) software has improved tremendously over the past several years. Two of the major commercial software packages for OCR are *Textbridge Pro 9* and *OmniPage Pro 10.* I used *Textbridge Pro* (an earlier version) during the process of creating the NIS data set. A variety of scanners have become available at relatively low cost during the last two years making OCR a reasonable option even for perpetually underfunded graduate students. One note of caution regarding scanner selection is in order though. A flat bed scanner requires attention every time a page is scanned, whereas a sheet fed scanner can be left running with up to the maximum number of pages the input tray can handle. I, therefore, strongly suggest a sheet fed scanner as a way of promoting the efficient use of time. Pages should be scanned in as line art and at an optimal resolution. I have used both a sheet fed scanner with a maximum resolution of 400 dpi and a flat bed scanner with a maximum resolution of 9600 dpi. I have found 400 dpi to be an excellent setting that provides enough information for the OCR package to be able to correctly identify even small fonts while not providing too much information–increasing the resolution beyond the optimal setting seems to result in additional phantom characters. Increasing the resolution also increases the size of the resulting file and requires additional computing power to handle. The outrageous advances in cheap computing power and storage make this less of an issue than even two years ago, but nonetheless, one should keep this in mind. Finally, OCR, although improved, is not perfect. *One absolutely must proofread the resulting text files.*

After spending months trying to prepare my data and making many mistakes, I have also found that I had much better success saving the file in blocks of no more than 20 pages. I suggest a 20 page limit because *Textbridge Pro* will occasionally crash when one is nearing the end of a long afternoon of text recognition causing one to lose all of the current work. Keep it short so that you can save frequently. Also, I strongly advise saving the file as text with no pictures directly from the OCR software.

Consistency in formatting is another important issue concerning the choice of source selection. Unless one is willing to retype and or reformat large portions of the source material, consistency in formatting is essential. Machine coding programs have format requirements that are unlikely to be met by the source one chooses. By choosing a source with a consistent format one can either write a program (e.g., in *Visual Basic*) or write a macro (e.g., in *Word Perfect* or *Word*) to reformat the source to meet the requirements of the coding program. I spent over 30 hours writing macros in *Word Perfect* to reformat the NIS raw data.[9] The specific formatting requirements will depend on the machine coding program that one selects.

---

9 A cautionary note of practical advice, *Word Perfect* will occasionally convert a hard-return code into a dormant hard-return code. If you write your macro to use hard-returns as anchors, you can easily miss observations during the reformatting process–I did, and its terribly frustrating to proofread more than a 1,000 pages of text for a second time.

4. *Choosing a coding program.* The Kansas Event Data System (KEDS), although showing age, remains the best known machine coding program for event data.[10] Two alternatives are under development. The first is Text Analysis by Augmented Replacement Instructions (TABARI), which is being developed by Philip Schrodt who also wrote KEDS, and the second is FRED, by Doug Bond who is part of the Protocol for the Assessment of Nonviolent Direct Action (PANDA) project. However, at this time KEDS appears to be the only readily available alternative.[11] KEDS was developed with the assistance of an National Science Foundation grant and is free. The program can be downloaded from the KEDS website: http://www.ukans.edu/~keds. One drawback to KEDS is that it only runs on a *Macintosh* platform. This can be a serious problem if one works primarily in a *Windows* environment. TABARI will run under LINUX which is a freely available operating system that can be downloaded from the Internet. LINUX can be dual booted with a *DOS/Windows* operating system making this a better option than buying a *Macintosh* and running the soon to be outdated KEDS.

The NIS data set was developed using KEDS. Since my primary system is a *Windows* based machine, using KEDS involved switching disks back and forth between the two machines. This creates a few additional problems. Whenever I would transfer a text file over to the *Macintosh*, a series of lead spaces would appear on every line. This could be remedied only by loading the file into a word processor and then resaving the file as a *Macintosh* text file. Since I had only bought an old *Macintosh*, it could not handle converting the entire data set at a single time. This required the additional step of breaking the data set into five files. The problem of lead spaces also occurred whenever I would transfer the data from the Macintosh to the *Windows* based machine. This is easily rectified by writing a macro that replaces the hard return at the end of the line and the space at the beginning of the new line with a hard return. Hopefully, all of this nonsense will be made obsolete by the appearance of TABARI.

5. *Formatting the Data.* As stated earlier, the specific formatting requirements are dependent upon the coding program. For KEDS, please refer to the KEDS manual for the specific formatting requirements. However, keep in mind the character per line limitation that the manual gives (96 characters). Saving files as ASCII text files often results in a hard return being added to the end of every line–i.e., all soft returns are converted to hard returns. Unless the source text has wide columns, one can normally proceed with the original columns. This makes reformatting much easier since one only has to put the header information for each entry.

When using KEDS one can choose a number of options, but one of the more interesting is whether to code lead sentences or entire stories. In the case of chronological data, this would mean that one would only be coding the first sentence to an entry or every sentence in an entry. Given that chronologies often report different events in the same paragraph, one might be better

---

10 Publications using KEDS coded data include Bond, Jenkins, Taylor, and Schock (1997); Gerner, Schrodt, Francisco, and Weddle (1994); Goldstein and Pevehouse (1997); Huxtable and Pevehouse (1996); Schrodt (1993; 1994; 1998); Schrodt and Gerner (1994; 1997; 1998); Schrodt and Savaiano (1997); and Schrodt, Davis, and Weddle (1994).

11 I have not been able to find any online mention of FRED at this time. The development of TABARI is documented on the KEDS website: http://www.ukans.edu/~keds/tabari.html.

served to code by sentence. This introduces additional complications though in the reformatting process since not only must one include a unique event code but also a sentence code. Once again, all of these things are best done automatically by writing a macro or a separate computer program. Editing by hand causes great fatigue and thus lends itself to errors. All of the reformatting work for the NIS project was accomplished using macros in *Word Perfect.*

6. *Choosing or Creating Actor and Verb Dictionaries.* Unless one is trying to create a historical event data set for an area that has already been studied using machine coded event data, developing an actor dictionary will be required. In terms of using KEDS and TABARI this requires developing a list of actors who you would like the program to recognize as either a source or a target of an action. The Kansas event data group has developed several utilities to assist in using KEDS; one of these programs helps identify proper names that occur in the raw data and actually counts the number of times each name is used. This is an excellent way to start an actor dictionary. Nonetheless, one must still sort through the program's output to select which names belong to actors, and in addition, one must also specify what association the actor will have. Another alternative that is more time consuming but worthwhile is to scan the raw data. This works well since one can more quickly identify the association of a particular actor by reading the name within the context of the original sentence or paragraph. The NIS actor dictionary has gone through five major revisions and now has several hundred actors associated with thirty-nine different political parties, paramilitaries, and social movements in Northern Ireland, Great Britain, the Republic of Ireland, as well as the United States, the European Union, the United Nations and Libya. In summary, actor dictionaries are tailored to the particular area and time period; however, verb dictionaries tend to be more independent of the area and time.

As a first step, choosing an existing verb dictionary is highly recommended since a typical verb dictionary may contain over 4,000 phrases. The KEDS verb dictionary can be downloaded with the program, and is an excellent base on which to build. The KEDS verb dictionary seems to have been optimized for coding events in the Middle East but can easily be applied to other areas as well. The NIS data set was built using a slightly modified version of the KEDS dictionary.[12] Another outstanding alternative is the PANDA project verb dictionary which was also developed to work with KEDS. This is available on the PANDA web page (http://data.fas.harvard.edu/cfia/pnscs). The PANDA dictionary is designed to identify a broader range of nonviolent direct action and protest behaviors than the KEDS dictionary. Both PANDA and KEDS verb dictionaries code data into a coding schemes that are derivatives of the original 63 WEIS categories (Bond, Jenkins, Taylor, and Schock 1997). PANDA currently uses 95 categories while the KEDS dictionary includes the 63 WEIS categories plus a number of additional codes resulting from an attempt to code events similar to those found in the PANDA categories.[13] The choice of a verb dictionary is significant since the coding scheme used to

---

12 In the process of optimizing the verb dictionary, I added around 100 phrases. Almost all of these were a result of my need to pick up on violence between social groups rather than states.

13 Surprises seem to always pop up when relying on other people's work, and this was a case in which that happened. Apparently someone added a number of additional categories based on the 22 broad WEIS categories. To my knowledge this information is not included in the manual or any of the other accompanying materials. I only found out where the new categories were coming from after I got unexpected codes and asked Philip Schrodt.

categorize events will undoubtedly “color” the results one gets from the raw data–this is analogous to Duffy’s (1994) world versions.

7. *Fine Tuning Coding*. Once the data has been appropriately formatted for the coding program and actor and verb dictionaries have been chosen, one can begin to fine tune the coding. This is accomplished through selecting different coding options and by modifying the actor and verb dictionaries. Major options include coding all events identified in a sentence or coding by the first event in a clause, determining what types of sentence structures to divert to a file due sentence complexity, and also setting what types of events are considered valid–i.e., is an identified source, target, and action required or only an action with either a source or target–etc. The complex sentence file is a key component in the process of fine tuning coding. Fine tuning involves coding the data a number of times. In between each run, output from the complex sentence file should be evaluated. By choosing to divert entries to the file if the program does not find a verb, source, target, or event allows one to make the necessary additions to the actor or verb dictionaries to allow the program to code all desired events. However, even with considerable work, not all of the entries that I thought the program should code were coded. Hopefully, TABARI will remedy some of these problems.[14] A list of the KEDS options used to code the NIS data set can be obtained upon request.

8. *Assessing Accuracy*. Machine coding is normally 100% reliable in that the same codes are associated with the save verb and actor each time, however, like human coding, it is not 100% accurate: some mistakes are made. The typical method of assessing the accuracy of machine coding (or human coding for that matter) is to assume that a coder (coders) has (have) privileged knowledge. This coder (coders) generates events from the raw data without referencing the machine coded material. The machine coded results are then compared with the results of the privileged human coder. Although one could evaluate the accuracy of each event, this would defeat the purpose of using a machine coding program. Therefore, one should randomly sample text entries to see whether or not the machine has “accurately” coded the entries.

Statistics on the relative success of KEDS in coding the three chronologies comprising the NIS data set are shown in Table 1. Coding has been optimized for the edited *Fortnight* chronology, 1970-1990, since it covers the largest amount of time. The grammatical structure of the entries also favors higher accuracy for the *Fortnight* abridged chronology. Its sentences are more concise and direct than those of the other two. The 82% coding accuracy for the *Fortnight* abridged chronology is equivalent to most human coded projects (Schrodt and Gerner 1994). Although the computer coded data from the other two chronologies is definitely more noisy than the *Fortnight* abridged chronology, this is still quite acceptable given the intended use of the data, which will be explained in the third part of this paper.

<< Table 1 About Here >>

---

14 Even though Schrodt insisted to me that the passive voice option that is set in the classes file was working, I was unable to get the feature to automatically code passive voice structures. Using the same options but doing machine assisted coding instead of straight machine coding without human intervention, the passive voice feature worked. Apparently, this feature caused other people problems since TABARI will have code written specifically for the problem.

9. *To Scale or Not to Scale, That is the Question.* Event data has frequently been scaled and aggregated. Other options certainly exist; however, scaling remains a major part of most event data analyses. Although the scaling of COPDAB events has changed through time--earlier Azar work uses a 13-point scale (Azar 1978)--the two ends of the spectrum have remained consistent:

> On the international scale, point 1 is the value given to the most cooperative event between two nations (e.g., Nation A and Nation B unite into one nation-state). Point 15 is the most conflictive event between two or more nations (e.g., total war). As for the domestic scale, point 1 is the most cooperative event (e.g., major governmental programs and policies to increase substantially socioeconomic freedom and equality), and on the other end of the domestic scale--point 15--is the most conflict and violent of physical events, namely civil war (Azar 1980, 149).

Efforts have also been made to scale the WEIS event data set (Vincent 1983; Goldstein 1992). However, McClelland himself remains quite skeptical about the scaling of WEIS nominal categories: ". . . I doubt that one can contend, seriously, that the ratios are correct. There is something bizarre in asserting that the act of sending home a diplomat for alleged spying (at 3.8) is 0.9 of a unit more conflict-laden than a public accusation that spying has been going on (2.9) but 0.1 less weighty than a threat delivered against the apparent offender" (1983, 174). McClelland also rejects scaling due to the context specific nature of conflict:

> The question of how much conflict is contained in a set of actions is approached, necessarily, in the environment of previous actions and in the context of the affective responses of participants and witnesses. *Conflict* is a second order meaning that is greatly influenced by judgements rendered about first order meanings. For instance, warnings which are so readily regarded as indicating hostility, tension, antagonism, opposition (or conflict) can be friendly, altruistic, concerned, integrative (or cooperative). The crucial factor is the situation and situational contexts are highly variable (1983, 174).

Therefore, for these reasons and also for the earlier stated philosophical difference in what events represent, McClelland rejects the cooperation/conflict approach.

Nonetheless, scaling of the WEIS nominal codes has become relatively standard practice with Goldstein's (1992) conflict-cooperation scale seemingly being the most accepted. Goldstein's scale is based on the weighting of events from -10.0 for the most hostile to 8.3 for the most cooperative. The weights were determined by averaging the weights assigned by a panel of eight scholars from the University of Southern California. The weights are shown in Table 2.

<< Table 2 About Here >>

Although the individual weights are quite questionable and are almost certainly affected by measurement error, aggregation of single events into longer aggregation periods containing a number of events improves the signal-to-noise ratio of the measure. The greater the number of events per aggregation period, the more accurate the measure is. Nonetheless, this accuracy is gained at the expense of information.[15] A very long aggregation period would lead to a highly accurate but probably useless figure for most questions and would require substantial supplemental contextual information since the ebb and flow of interaction would be lost. Consequently, the length of the aggregation period should depend on the temporal density of events and the research question. Some research areas, such as crisis, require short aggregation periods solely due to the natural time of the process under study. The NIS data exists in several forms: unscaled, scaled, disaggregated, and aggregated in periods of one day to six weeks.

One problem with aggregating the data is that effects of aggregation are not always apparent. The mean of the aggregated data scales directly in relation to the aggregation period. However, when one looks at the variance and skewness of the data, the second and third moments about the mean respectively, the results are not directly predictable and depend on the sequencing of events. See Figure 2. Thus, care should be taken in choosing an aggregation period.

<< Figure 2 About Here >>

The Kansas event data group provides a program, KEDS_COUNT for scaling and aggregating the KEDS output. However, some versions of the program drop observations. In the case of the NIS data, approximately 50% of the original events were unintentionally dropped. As a result, I have written a program in *Visual Basic* that is available upon request. The program currently creates a systemic total rather than dyadic totals for an aggregation period, but can be modified with relative ease to aggregate on the basis of dyads. More recent versions of the Kansas program may have fixed the problem of dropping events.

Scaling and aggregation introduces additional measurement error into the data, and a number of authors are turning (or returning) to analyses that do not require that event data be treated as interval level data. Schrodt's (1997a; 1997b) recent work on Hidden Markov Modeling is an excellent example. His analysis returns to original efforts by McClelland to identify patterns in the sequencing of the event data nominal categories. Alternatively, Duffy (1994) suggests a computational hermeneutic approach. Thus, the question of to scale or not to scale is dependent upon the research question, which should itself be shaped by the interplay between theory and research design.

USING THE NIS DATA SET

The NIS data set was originally developed as part of my dissertation research. In my dissertation, "Crisis, Equilibrium, and Protracted Social Conflict: The 'Strange Attractiveness' of Protracted Social Conflict in Northern Ireland," I answer Charles McClelland's 1972 question, "Is it possible to identify crises by means other than the general feel of the situation or the judgements of contemporary observers and participants?" within the specific context of protracted social conflict as conceptualized by Edward Azar. In doing so, I examine the relative

15 Information theory defines the amount information in a signal as the amount of uncertainty. Aggregation reduces the degrees of freedom and thus uncertainty and the amount of information in a signal.

strengths of the implicit theory of crisis–the commonly accepted belief that crises represent qualitatively distinct periods that are therefore worthy of scholarly attention–and prospect theory as applied crisis in protracted social conflict. Three competing philosophical approaches along with their respective methods–determinism and dynamic systems analysis, probabilism and statistical analysis, and historical analogy and hidden Markov modeling–are used to inductively identify crisis periods on the basis of the original systemic level NIS event data set covering the years 1968 - 1990 and 1992 - 1996. I also propose an equilibrium theory of normal relations as a robust alternative to Azar's normal relations range. The results of the analyses indicate that the implicit theory of crisis is generally supported, and in relation to crisis in protracted social conflict, it shows greater predictive and explanatory power than prospect theory with its emphasis on issue framing. In sum, on the basis of the analyses of the NIS data, crisis in protracted social conflict does appear to be inductively identifiable through examining systemic level cooperation and conflict flows.

The NIS data has also been used to develop a measure of conflict scope in the Northern Ireland domestic political conflict. Gurr et al. (1978) define conflict scope as citizen participation in open confrontation (riots, marches, violence, etc.), measured in "man-days" per 100,000 population. In my recent study of the impact of domestic political conflict in Northern Ireland on overall social well-being for the society, I use the number of events per year as identified by the NIS data set showing citizen participation in open confrontation divided by that current years population in Northern Ireland. Although this measure is not strictly the same as the original Gurr et al. (1978) variable, the number of incidents appears to be a workable substitute given that little agreement exists about numbers of participants for events in Northern Ireland. The study uses only a small number of the original 63 nominal categories found in the data; however, the availability of the data makes the study possible. The results of the study suggest that conflict scope and intensity are curvilinearly related to social well-being. The functional form of the relationship between conflict scope and social well-being is concave-up, while that of conflict intensity is concave-down. These results are similar to those found in Dixon and Moon (1989), but their study was unable to uncover the hypothesized curvilinear relationships.

These two research projects are examples of the type of research that one can conduct using historical event data sets. The combination of machine coding and optical character recognition software makes the construction of new historical event data sets feasible both in terms of time and money. However, if such research is to become more common, a conscious community of scholars engaged in the development of historical event data sets should be developed. The sharing of actor and verb dictionaries, macros, other utility programs, and advice on OCR is an important step to insuring that this type of research can be conducted on a wider basis. In sum, the resurgence of current event data research should be met with a similar surge in historical event data research.

REFERENCES

Alker, Hayward R., Jr. 1993. "Making Peaceful Sense of the News: Institutionalizing International Conflict-Management Event Reporting Using Frame-Based Interpretive Routines." In *International Event-Data Developments: DDIR Phase II*, ed. R.L. Merritt, R.G. Muncaster, and D.A. Zinnes. Ann Arbor: University of Michigan Press.

Azar, Edward. 1972. “Conflict Escalation and Conflict Reduction in an International Crisis: Suez, 1956.” *Journal of Conflict Resolution* 16 (June): 183-201.

Azar, Edward. 1978. “An Early Warning Model of International Hostilities.” In *Forecasting in International Relations: Theory, Methods, Problems, Prospects,* ed. N. Hourchi and T. Robinson. San Francisco: WH Freeman.

Azar, Edward. 1980. “The Conflict and Peace Data Bank (COPDAB) Project.” *Journal of Conflict Resolution* 24 (March): 143-252.

Azar, Edward, James P. Bennet, and Thomas J. Sloan. 1974. “Steps Toward Forecasting International Interactions.” *The Papers of the Peace Science Society (International)* 23: 27-67.

Bell, Robert, Robert Johnstone, and Robin Wilson, eds. 1991. *Troubled Times: Fortnight Magazine and the Troubles in Northern Ireland 1970-91.* Belfast, Ire: Blackstaff Press.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor, and Kurt Schock. 1997. “Mapping Mass Political Conflict and Civil Society: The Automated Development of Event Data.” *Journal of Conflict Resolution* 41.4: 553-579.

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

Davies, John L., and Chad K. McDaniel. 1994. “A New Generation of International Event-Data.” *International Interactions* 20.1-2: 55-78.

Deutsch, Richard, and Vivien Magowan. 1973. *Northern Ireland 1968-73: A Chronology of Events Volume 1 1968-71.* Belfast, Ire.: Blackstaff Press.

Dixon, William, and Bruce Moon. (1989) Domestic Political Conflict and Basic Needs Outcomes: An Empirical Assessment. *Comparative Political Studies* 22 (July 1989): 178- 198.

Duffy, Gavan. 1994. “Events and Versions: Reconstructing Event Data Analysis.” *International Interactions* 20.1-2: 147-167.

Gerner, Deborah J., Philip A. Schrodt, Ronald A. Francisco, and Judith L. Weddle. 1994. “Machine Coding of Events Data Using Regional and International Sources.” *International Studies Quarterly* 38 (March): 91-119.

Goldstein, Joshua S. 1992. “A Conflict-Cooperation Scale for WEIS Events Data.” *Journal of Conflict Resolution* 36 (June): 369-85.

Goldstein, Joshua S., and Jon C. Pevehouse. 1997. Reciprocity, Bullying and International Cooperation: A Time-Series Analysis of the Bosnia Conflict.” *American Political Science Review* 91.3: 515-530.

Gurr, Ted, et al. 1978. *Comparative Studies of Political Conflict and Change: Cross National Datasets*. Ann Arbor, MI: ICPSR.

Gurr, Ted R. 1975. *Polimetrics.* Englewood Cliffs, N.J.: Prentice Hall.

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

Huxtable, Phillip A. and Joc C. Pevehouse. 1996. “Potential Validity Problems in Events Data Collection.” *International Studies Notes* 21.2: 8-19.

Lave, Charles A., and J.G. March, eds. 1975. *An Introduction to Models in the Social Sciences.* New York: Harper and Row.

Lebovic, James. 1994. “Before the Storm: Momentum and the Onset of the Gulf War.” *International Studies Quarterly* 38: 447-474.

Mallery, John. 1994. "Beyond Correlation: Bringing Artificial Intelligence to Events Data." *International Interactions* 20.1-3" 101-145.

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 (June): 169-77.

McClelland, Charles A., and Gary D. Hoggard. 1969. "Conflict Patterns in the Interactions Among Nations." In *International Politics and Foreign Policy: A Reader in Research and Theory,* ed. James Rosenau. New York: Free Press.

Merritt, R.L, R.G. Muncaster, and D.A. Zinnes, eds. 1993. *International Event-Data Developments: DDIR Phase II*. Ann Arbor: University of Michigan Press.

Most, Benjamin, and Harvey Starr. 1989. *Inquiry, Logic and International Politics.* Columbia, SC: University of South Carolina Press.

Schrodt, Philip A. 1993. "The Machine Coding of Events Data." In *International Event-Data Developments: DDIR Phase II*, ed. R.L. Merritt, R.G. Muncaster, and D.A. Zinnes. Ann Arbor: University of Michigan Press.

Schrodt, Philip A. 1994. "The Statistical Characteristics of Event Data." *International Interactions* 20.1-2: 35 - 53.

Schrodt, Philip A. 1997a. "Pattern Recognition of International Crises Using Hidden Markov Models." Paper Presented at the Annual Meeting of the International Studies Association, Toronto.

Schrodt, Philip A. 1997b. "Early Warning of Conflict in Southern Lebanon using Hidden Markov Models." Paper Presented at the Annual Meeting of the American Political Science Association, Washington D.C.

Schrodt, Philip A. 1998. *KEDS: Kansas Event Data System.* Version 0.9B7.

Schrodt, Philip A., and Deborah J. Gerner. 1994. "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982-92." *American Journal of Political Science* 38 (August): 825-54.

Schrodt, Philip A., and Deborah J. Gerner. 1997. "Empirical Indicators of Crisis Phase in the Middle East, 1979-1995." *Journal of Conflict Resolution* 41 (August): 529-552.

Schrodt, Philip A., and Deborah J. Gerner. 1998. "Cluster Analysis as an Early Warning Technique for the Middle East." In *Preventive Measures: Building risk Assessment and Crisis Early Warning Systems,* ed. John L. Davies and Ted Robert Gurr. Lanahm, MD: Rowman and Littlefield.

Schrodt, Philip A., and Scott Savaiano. 1997. "Environmental Change and Conflict: analyzing the Ethipian Famine of 1984-85." In *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts,* ed. Carl W. Roberts. Hillsdale, NJ: Lawrence Erlbaum Associates.

Schrodt, Philip A., Shannon G. Davis, and Judith L. Weddle. 1994. "Political Science: KEDS-A Program for the Machine Coding of Event Data." *Social Science Computer Review* 12.3: 561-588.

Singer, J.D. and M. Small. 1972. *The Wages of War, 1816-1965: A Statistical Handbook.* New York: Wiley.

Vincent, Jack. E. 1983. "WEIS vs. COPDAD: Correspondence Problems." *International Studies Quarterly* 27 (June): 161-68.

Woolley, John T. 2000. "Using Media-Based Data in Studies of Politics." *American Journal of Political Science* 44 (January): 156-173.

TABLE 1: Goldstein Scale Weights

Event Type	Weight	
223 Military attack; clash; assault	-10.0	
211 Seize position or possessions	- 9.2	
222 Nonmilitary destruction/injury	- 8.7	
182 Armed force mobilization, exercise, display; military buildup	- 7.6	
195 Break diplomatic relations	- 7.0	
173 Threat with force specified	- 7.0	
174 Ultimatum; threat with negative sanction and time limit	- 6.9	
172 Threat with specific negative nonmilitary sanction	- 5.8	
193 Reduce or cut off aid or assistance; act to punish/deprive	- 5.6	
181 Nonmilitary demonstration, walk out on	- 5.2	
201 Order person or personnel out of country		- 5.0
202 Expel organization or group	- 4.9	
150 Issue order or command, insist, demand compliance		- 4.9
171 Threat without specific negative sanction stated	- 4.4	
212 Detain or arrest person(s)	- 4.4	
191 Reduce routine international activity; recall officials		- 4.1
112 Refuse; oppose; refuse to allow	- 4.0	
111 Turn down proposal; reject protest, demand, threat	- 4.0	
194 Halt negotiation	- 3.8	
122 Denounce; denigrate; abuse	- 3.4	
160 Give warning	- 3.0	
132 Issue formal complaint or protest	- 2.4	
121 Charge; criticize; blame; disapprove	- 2.2	
191 Cancel or postpone planned event	- 2.2	
131 Make complaint (not formal)	- 1.9	
063 Grant asylum	- 1.1	
142 Deny an attributed policy, action, role or position	- 1.1	
141 Deny an accusation	- 0.9	
023 Comment on situation	- 0.2	
102 Urge or suggest action or policy	- 0.1	
021 Explicit decline to comment	- 0.1	
094 Request action; call for	- 0.1	
025 Explain or state policy; state future position	0.0	
091 Ask for information	0.1	
011 Surrender, yield to order, submit to arrest	0.6	
012 Yield position; retreat; evacuate	0.6	
031 Meet with; send note	1.0	
095 Entreat; plead; appeal to; beg	1.2	
101 Offer proposal	1.5	
061 Express regret; apologize	1.8	
032 Visit; go to	1.9	
066 Release and/or return persons or property	1.9	
013 Admit wrongdoing; apologize, retract statement	2.0	
062 Give state invitation	2.5	
054 Assure; reassure	2.8	
033 Receive visit; host	2.8	
065 Suspend sanctions; end punishment; call truce	2.9	
082 Agree to future action or procedure, to meet, or to negotiate	3.0	
092 Ask for policy assistance	3.4	
093 Ask for material assistance	3.4	
041 Praise, hail, applaud, extend condolences		3.4
042 Endorse other’s policy or position; give verbal support	3.6	
053 Promise other future support	4.5	
051 Promise own policy support	4.5	
052 Promise material support	5.2	
064 Grant privilege; diplomatic recognition; de facto relations	5.4	
073 Give other assistance	6.5	
081 Make substantive agreement	6.5	
071 Extend economic aid; give, buy, sell, loan, borrow	7.4	
072 Extend military assistance		

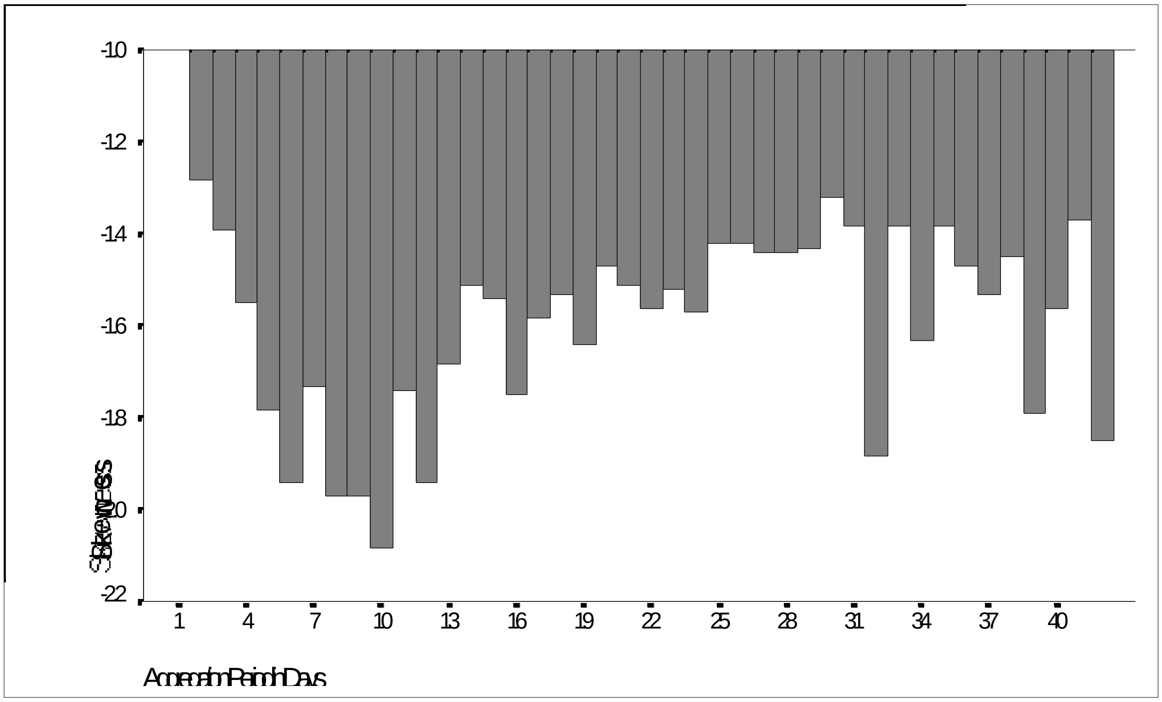

TABLE 2: KEDS Coding Accuracy

	Number of Entries	Number of Events[@]	Events Coded as Percentage of Entries	Percentage of Correctly Coded Events[*]
Deutsch and Magowen 1968 - 1970	2,691	2,240	83.2%	67% Å 9.3% 57.7% < ÿ < 76.3%
Edited *Fortnight* Chronology 1970 - 1990	4,616	4,572	99.04%	82% Å 7.5% 74.5% < ÿ < 89.5%
Unedited *Fortnight* Chronology 1992 - 1996	3,686	3,766	102.1%	73% Å 8.7% 64.3% < ÿ < 81.7%

[*]95% confidence interval, n=100
[@]Single chronological entries may generate more than one event. For example, if two parties meet, two events are generated, one for each party. This accounts for the surprising occurrence of more events being coded than there are chronological entries for the *Fortnight* chronology.