The Kansas Event Data System: A Beginner's Guide with an Application to the Study of Media Fatigue in the Palestinian *Intifada*

Deborah J. Gerner and Philip A. Schrodt
Department of Political Science
University of Kansas
Lawrence, KS 66045 USA
phone: 913-864-3523 fax: 913-864-5700
p-schrodt@ukans.edu d-gerner@ukans.edu

Poster session presentation at the American Political Science Association
San Francisco, 28 August - 1 September 1996

This paper, the KEDS program, data sets and other information are available at the KEDS web site:
http:// raven.ukans.cc.edu / ~keds

ABSTRACT

This paper provides a general introduction to using KEDS, the Kansas Event Data System. KEDS is a Macintosh-based machine coding system for generating event data using pattern recognition and simple linguistic parsing. The system codes from machine-readable text describing international events; the NEXIS data service, optical character recognition and CD-ROM can provide such texts. The paper is targeted at researchers who are considering using KEDS to generate event data: it describes the overall process of generating and analyzing event data, as well as providing a FAQ (Frequently Asked Questions) section and an annotated bibliography of sources of information on event data.

We illustrate the use of KEDS in political science research by examining the issue of "media fatigue": how does the number and type of events reported in public sources change as a conflict evolves? Using the first three years of the Palestinian *intifada* as a case study, we compare a machine-coded, Reuters-based event data set, a human-coded data set based on the *New York Times* and an independently-collected data set reporting levels of lethal violence during this period. As predicted by the media fatigue hypothesis, we find that the correlation between the three sources declines over time, and is generally proportional to the level of interest that the international media is showing to this region. The correlation between Reuters and the *New York Times* changes in a pattern that is similar to the correlation between these sources and the independent data source, suggesting that the correlation between the two news sources could be used as an indicator of the level of error in an event data set caused by media fatigue.

Development of KEDS was funded by the National Science Foundation through Grants SES89-10738, SBR-9410023 and SES90-25130 (Data Development in International Relations Project) and the University of Kansas General Research Fund Grant 3500-X0-0038. Brad Bennett, Doug Bond, Joe Bond, Shannon Davis, Ronald Francisco, Deborah Gerner, Phillip Huxtable, Tony Nownes, Jon Pevehouse, Julia Pitner, Uwe Reising, Scott Savaiano and Judy Weddle have served the difficult role of alpha- and beta-testers of KEDS

INTRODUCTION

The Kansas Event Data System (KEDS) is a system for the machine coding of international event data based on pattern recognition and sparse parsing of natural language reports. It is designed to work with short news summaries such as those found in wire service reports. To date, KEDS has primarily been used to code WEIS events (McClelland 1976) from the Reuters news service lead sentences but in principle it can be used for other event coding schemes and news sources.

Historically, event data have usually been hand-coded by legions of bored undergraduates flipping through copies of the *New York Times*. Machine coding provides two advantages over these traditional methods:

- Coding can be done more *quickly* by machine than by hand; in particular the coding of a large machine-readable data set by a single researcher is feasible, and a data set can be maintained in real-time with only a modest level of effort;
- Machine coding rules are applied with complete *consistency* and are not subject to inter-coder disparities caused by fatigue, differing interpretations of the coding rules or biases related to the texts being coded;

The disadvantage of machine coding is that it cannot deal with sentences having a complex syntax and it deals with sentences in isolation rather than in context.

KEDS can be used for either machine-assisted coding or fully automated coding. Coded events can be manually edited on the screen before they are written to a file, and the program has a "complexity detector" that can divert linguistically complex sentences—for example those containing a large number of verbs or subordinate clauses—to a separate file for later human coding.

How KEDS Works

KEDS combines simple syntactic analysis and pattern recognition to do its coding. Three types of information are used:

Actors: These are proper nouns that identify the political actors recognized by the system; Table 1a shows an example of the beginning of an actors dictionary;[1]

Verbs: Event data categories are primarily distinguished by the actions that one actor takes toward another, so the verb is usually the most important part of a sentence for determining the event code. Table 1b shows an example of the beginning of a verbs list.

Phrases: Phrases are used to distinguish different meanings of a verb—for example PROMISED TO SEND TROOPS versus PROMISED TO CONSIDER PROPOSAL—and to provide syntactic information on the location of the source and target within the sentence. The strings beginning with "-" in Table 1b are verb phrases, where "*" shows the location of the verb in the phrase.

KEDS relies on *sparse parsing* of sentences—primarily identifying proper nouns (which may be compound), verbs and direct objects within a verb phrase — rather than using full syntactical analysis. As a consequence KEDS will make errors on complex sentences or sentences using unusual grammatical constructions, but it requires less information to deal with the sentence structures that are most commonly encountered in news articles and has proven quite robust in correctly interpreting those types of sentences.

---

[1] The information following the semicolon identifies the coder and the date when the information was added to the dictionary.

Table 1a. Actors Dictionary

ABADAN [IRN] ;JON 5/24/95
ABU_DHABI [UAE] ;JW 11/7/91
ABU_NIDAL [PLOFA] ;jw 9/12/91
ABU_SHARIF [PLO] ; PLO AIDE ;tony 3/13/91
AFGHAN [AFG] ;jw 11/7/91
AFRICAN_NATIONAL_CONGRESS [SAFOP] ;jw 11/22/91
AFRICA [AFR] ;jw 11/20/91
AL-HARIRI [LEB] ;JON 4/24/95
AL-KASM [SYR] ;JON 5/15/95
ALBANIA [ALB] ;jw 11/12/91
ALEPPO [SYR] ;JON 5/15/95
ALEXANDER_HAIG [USA] ;JON 5/29/95
ALGERIA [ALG] ;jw 9/14/91
ALI_ABDULLAH_SALEH [YEMNO] ;shan 2/23/93
AL_GORE [USA] ;JON 5/7/95
AMAL [LEBAM] ;jw 9/9/91

Table 1b. Verbs Dictionary

ABDUCT [213] ;shan 7/24/93
- * OF + IN $ [213] ;JON 5/26/95
ABIDE [080] ;shan 2/20/93
- WOULD NOT * [111]
- FAIL * [111] ;JON 5/8/95
ABSTAIN [111] ;jw 10/25/91
ACCEPT [081:081] ;jw 11/14/91
- WILL DECIDE WHETHER TO * [025]
- SAID WOULD NEVER * [112]
- NOT * PROPOSAL [111]
- * FORMULATION [042]
- * CREDENTIALS [064]
- * INVITATION [082]
- FORCE TO * [150]
- DOES NOT * [111]
ACCUS [121] ;jw 11/14/91
- * SUPPORT TERROR [121] ;JON 5/5/95
- * AGGRESSION [121] ;JON 5/8/95
- * PRISONER [---]
ACKNOWLEDGE [023]
ADD [---]
- * WANTED [102] ;tony 4/19/91
- * VOICE [023]
ADOPT [081:081] ;PAS 9/24/91
- * RESOLUTION [102]
- * PROGRAM [084]
AFFIRM [042]

Advantages of machine coding

We originally became involved with machine coding because it is dramatically faster and less expensive than human coding. Once a researcher has established vocabulary lists of actors and verb phrases, the only cost involved in generating event data is the acquisition of machine-readable news reports. (These are increasingly available from CD-ROM and electronic networks.) Furthermore, a coding system developed at one institution can be used by other researchers through the sharing of vocabulary lists and coding software; this has been part of our collaboration with the PANDA project (see below).

In working with KEDS, we discovered two additional advantages to machine coding. First, it is free of non-reproducible coding biases. Human coding is subject to systematic biases because of assumptions made implicitly by the coders. For example, Laurance (1990) notes that even expert coders at the U.S. Naval Postgraduate School tended to over-estimate the military capability of China because they knew China to be a large Communist country. Because most human event coding is done part-time by students, coder biases are difficult to control. In contrast, in machine-coding the words describing an activity will receive the same code irrespective of the actors or time period involved. Sparse parsing also tends to make *random* coding errors that a good statistical analysis can rectify, whereas the systematic errors introduced by human coders are much more difficult to correct. Any biases embedded in the machine coding system are preserved *explicitly* in its vocabulary; there is no such record in human coding.

Second, it is much easier to experiment with alternative coding rules using machine coding. The COPDAB (Azar 1982) and WEIS (McClelland 1976) event coding schemes are very general and for the most part presuppose a Cold War, Westphalian-Clausewitzian conflict framework. This weakens the value of WEIS and COPDAB data when dealing with post-Cold War phenomena such as ethnic conflict, low-intensity conflict, and multilateral intervention. Using a machine-coding system, even a very large data collection such as our 1979-1996 Arab-Israeli conflict data set (80,000 events) can be completely recoded in a couple hours. This is impossible with human coded data, which has severely restricted experimentation with new coding schemes.

History

The development of KEDS began around 1990 as part of the National Science Foundation's "Data Development in International Relations" project (Merritt, Muncaster & Zinnes 1994). While the DDIR work included some experimentation with German-language sources and foreign policy chronologies (Gerner et al. 1994), most of our experimentation at Kansas has been done in the English-language with WEIS coding and interactions reported by Reuters for the Middle East. We have primarily been coding "lead" sentences: the first sentences of the article which usually succinctly summarizes the main political event being reported in the story. We are currently maintaining an event data set covering the Levant and Gulf areas of the Middle East that covers 1979 to the present.

The other major project using the KEDS program is the Protocol for the Assessment of Nonviolent Direct Action (PANDA) at the Program on Nonviolent Sanctions and Cultural Survival at the Center for International Affairs at Harvard University (Bond, Bennett & Vogele 1994; web site: hdc-www.harvard.edu / cfia / pnscs / panda1.htm). This project uses KEDS to code a superset of the WEIS categories (about 100 categories versus the 63 categories in WEIS) that provide far more detail on nonviolent events, substate actors and internal interactions such as strikes and protests than are provided in WEIS. The PANDA scheme codes several contextual variables in addition to the standard date-source-event-target variables of event data. Reuters reports dealing with the entire world have been coded for 1985-1995; the resulting data set contains about 500,000 events.

In addition to these large projects, several additional KEDS dictionaries have been developed by our graduate students and others:

- Behavioral Correlates of War (BCOW; Leng, 1987) coding system of the Middle East (Jon Pevehouse)
- WEIS coding for West Africa, using full-story coding (Phillip Huxtable)
- WEIS coding for the conflict in the former Yugoslavia (Jon Pevehouse and Joshua Goldstein)
- Machine-assisted coding of political demonstrations in Europe (Ronald Francisco and Uwe Reising)

System Accuracy

The accuracy of KEDS depends heavily on the source text, the event coding scheme and the type of event being coded. We have reported a variety of different reliability checks in papers and articles (e.g. Gerner et al. 1992, Schrodt & Gerner 1994); other tests of the KEDS system are found in Huxtable and Pevehouse (1996) and Bond et al. (1996).

In the data set we are developing for the Middle East—Reuters lead sentences and the WEIS coding scheme—KEDS assigns the same code as a single human coder in about 75% to 85% of the cases. Approximately 10% of the Reuters leads have a syntactic structure that is too complicated or too idiosyncratic for KEDS to handle properly, although some of the residual coding disagreement comes from ambiguities in the WEIS coding categories themselves.[2] In an experiment where dictionaries were optimized for the coding of a single day of Reuters leads, the PANDA project—using a coding scheme substantially more detailed than WEIS—achieved a 91.7% machine coding accuracy; this probably represents the upper limit of accuracy for Reuters leads and a program using KEDS's sparse parsing approach (Bond, Bennett & Vogele 1994:9). This level of coding accuracy is comparable to that achieved in event data projects using human coders: Burgess & Lawton (1972:58) report a mean intercoder reliability of 82% for eight projects where that statistic is known.

Schrodt & Gerner (1994) assess the face validity of KEDS-generated data for the Middle East, 1982-1993; the time series produced by the program correspond closely to the patterns expected from narrative accounts of the interactions between the actors. In these papers, the KEDS data were also compared to a human-coded WEIS data set for 1982-1991 (Tomlinson 1993). For almost all dyads, there was a statistically significant correlation between the number of events reported by the two series, as well as the number of cooperative events. Event scores aggregated using the Goldstein (1992) scale and number of conflictual events showed a statistically significant correlation in about half of the dyads. Many of the differences between the two series appear to be due to the higher density of events in KEDS compared to the *New York Times*-based WEIS: the Reuters series contained, on average, three times as many events as WEIS. In this article we also duplicated two statistical studies—one involving cross-correlation and the other spectral analysis—using the machine-coded and human-coded data sets and found that they produced generally comparable results, albeit with some idiosyncratic differences in specific dyads.

---

[2] Examples of sentences that are too complex to code include the following:

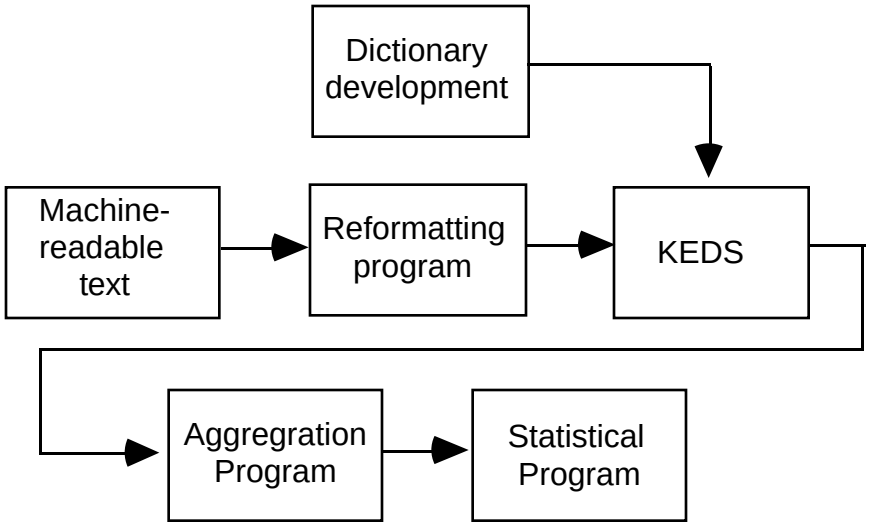
> The United States on Friday dismissed Israel's apparent rejection of an Egyptian plan for talks with the Palestinians as 'parliamentary maneuvering' and said the door was not closed to peace.
>
> Resumption of ties between Egypt and Syria may spur reconciliation between Iraq and Syria, and Syria and the PLO, the Qatari newspaper *al-Raya* said on Friday.

KEDS now includes a "complexity filter" that prevents the coding of sentences that are likely to be incorrectly interpreted by a sparse parser. For example, we do not code events in our Middle East data set if a sentences contains six or more verbs or if it has no actor before the verb. Our most recent analyses (Schrodt and Gerner 1996) suggest that this results in a data set with less noise than the data sets we produced without the complexity filter.

GENERATING DATA WITH KEDS

Because both the natural language text *input* and the event data *output* of KEDS are uncommon in the standard statistical data processing environment, working with KEDS requires a few more steps than, say, pulling variables off the Euro-Barometer surveys. Figure 1 below shows the process involved in going from machine-readable text to data that can be analyzed with a statistical program.

Figure 1.



In the section below, we will discuss the content of each of these steps in the data-generation process.

STEP 1: Locate and reformat a set of machine-readable texts

The very first step in doing research with KEDS is finding a source of machine-readable text. This will usually come from an on-line data service such as LEXIS/NEXIS or CD-ROM.[3]

In all likelihood, the original text will not be in the format used by KEDS. For example, the text provided by the NEXIS data service that we've been using originally looks like

[3] In Gerner et al. (1994) we discuss our experiments with using a scanner and optical character recognition (OCR) to generate machine-readable text from printed documents. We had mixed results with this but OCR software was relatively primitive at that point and the technology has apparently improved substantially in recent years.

```
] ] ]
]                          LEVEL 1 - 5 OF 914 STORIES
] ]             Proprietary to the United Press International 1981
] ]                     <December> 29, 1981, Tuesday, PM cycle
] ]HEADLINE: World News Summary
] ] BODY:
]   Sen. Charles Percy, R-Ill., chairman of the Senate Foreign
]Relations Committee, asked his Israeli hosts to avoid the annexed
]
^
^
]Golan Heights in a helicopter inspection tour today of the tense
]Israeli-<Lebanese>frontier.
]
```

This needs to be converted to the KEDS input format

```
811229UPI001
Sen. Charles Percy, R-Ill., chairman of the Senate Foreign
Relations Committee, asked his Israeli hosts to avoid the annexed
Golan Heights in a helicopter inspection tour today of the tense
Israeli-Lebanese frontier.
```

We've developed a number of reformatting programs that remove the irrelevant information found in the NEXIS download and reformat the text; some of these programs and their Pascal or C source code can be found at the KEDS web site.

If your original text is not in exactly the same format as that we found in NEXIS, you may need to write a filter or modify one of ours. Because machine-readable data are generally consistently formatted, this is usually not very difficult provided you know (or know someone who knows) a programming language such as BASIC, C, Pascal or PERL, but such a filter is a necessary step before you can start using KEDS. If you aren't able to get a filter/reformatting program written in a programming language, the macro language in Microsoft's *Word* program provides another possibility for reformatting; this has been used by some projects.

STEP 2: Develop the initial coding dictionaries

KEDS uses large dictionaries of proper nouns and verb phrases to code the actors and events it finds in the source text. If you intend to code political events, you would probably find it easier to modify the dictionaries that have been developed by other projects rather than starting with a new dictionary. These dictionaries have already identified most of the English vocabulary used in Reuters, so even if you expect to substantially change the coding scheme, you will know what types of phrases to expect. The KEDS (WEIS, Middle East) and Pevehouse (BCOW, Middle East) dictionaries are available from our web site and are archived at the ICPSR; the PANDA dictionaries are available from the PANDA project (contact: dbond@cfia.harvard.edu). All three of these dictionaries were based on coding Reuters lead sentences.

STEP 3: Fine-tune the dictionaries

With the initial dictionaries incorporated into the system, the next step is fine-tuning—"tweaking"—the phrases to work correctly with your data and coding scheme. This is done by going through a large number of texts and modifying the vocabulary as needed; this process will also provide an indication of the accuracy of the system. Most vocabulary modifications involve the addition of specific individual actors (e.g. political leaders; geographical place names) and the addition of verb phrases describing behaviors specific to the problem being studied.

While fine-tuning the dictionaries you might also look at some of the advanced features of KEDS, such as the use of substitution rules, word classes, the complexity filter, and additional coding features such as issues and content analysis counts. Grammatical transformation rules can sometimes provide general solutions to problems that would otherwise require a large number of specific phrases.[4] The additional coding features allow information to be extracted from a sentence beyond the basic *source-event-target* structure of event data when this is required by the coding scheme.

Resist the temptation to tweak the vocabulary indefinitely! Tweaking should focus on finding *general* patterns that occur on multiple occasions in the text, not on expanding the list of phrases to cover every possible contingency. Coding errors are only *one source* of noise in event data: the source text is already an incomplete and biased record of the underlying events; the event coding scheme may be incorrectly aggregating categories of events; and the statistical models for which the data will eventually be used in capture only some of the possible forms of the political relationships.

Even if your coding was somehow "perfect", you are still dealing with a noisy process. Successfully identifying relationships amid that noise comes at the *analytical* stage of the project—both in the development of the coding scheme and the statistical analysis—not in the coding stages. Know when the coding accuracy is "good enough", and don't fall into the trap of doing beer-budget analysis on champagne-budget data. If you can't cope with the fact that probably 15% of your data are erroneously coded—and this is true whether you are using human or machine coding—you shouldn't be doing event data analysis. End of sermon.

KEDS recognizes the following types of words, which are called "classes" in the KEDS system. Some of these correspond to conventional parts of speech (e.g. pronouns, conjunctions), while others are more specific (e.g. actors and agents).

Class	Content
actor	Nouns in the *.Actors* dictionary
agents	Improper nouns such as POLICE, DEMONSTRATORS
im.actr	Implicit actor: an agent that is treated as an actor
verb	Verbs in the *.Verbs* dictionary (but not words inside verb phrases)
conj	Conjunctions: AND_ and BUT_
pronoun	Pronouns: HE_ SHE_ IT_ THEY_, THEM_ and ITS_
prep	Prepositions: these are used only in the optional **Places** facility that attempts to identify the location of an event
comma	Commas
stop	Articles: A_ AN_ THE_ (these are discarded)
null	Null-coded actors, null-coded verbs without phrases, and any word not in the dictionary

[4] For example the reversal of source and target in English passive voice—"ISRAEL WAS CRITICIZED BY SYRIA" versus "SYRIA CRITICIZED ISRAEL"—and the use of markers such as AS, WHILE, and WHEN as conjunctions.

The standard KEDS parsing does the following:

- Identifies compound actors
- Reduces titles to a single actor reference
- Identifies compound clauses within a sentence
- Locates the references of pronouns
- Discards stop words and "null-coded" words[5]
- Eliminates common-delimited subordinate clauses

Additional parsing features can be activated by using commands in the *.Options* file:

- Association of agents with actors
- Evaluation of prepositional phrases for determining location

STEP 4: Autocode the entire data set

Unless you intend to use KEDS for machine assisted-coding of an entire dataset, the data should be autocoded after the accuracy of the dictionaries has reached an appropriate level. Autocoding will ensure that the coding rules have been consistently applied across the entire data set, rather than having the part of the data coded by hand, and the remainder machine coded. Data which has been partially hand-coded and partially machine-coded is likely to cause statistical artifacts in time series analysis, for example by making some types of behavior more common in the human-coded parts of the data than in the machine-coded parts. Autocoding also insures that your coding can be replicated by later researchers, as well as extended at a later date.

If you cannot get the accuracy level of KEDS to an acceptable level, you still may be able to use the program for routine coding by using KEDS's complexity filter. This will systematically divert to a separate file any texts that appear too complex to machine code, for example those containing an excessive number of verbs or actors, or containing ambiguous words such as ATTACK or GEORGIA. The complex texts can be processed using machine-assisted coding, and then the two sets of event data can be merged. Coding with the complexity filter is not completely replicable—and the complexity filter will not catch all sentences that might be coded incorrectly—but it is more efficient and replicable than all-human coding, while still remaining more accurate than all-machine coding.

STEP 5: Aggregate the data for statistical analysis

KEDS produces standard event data of the form[6]

<date> <source code> <event code> <target code> <auxiliary codes>

Because event data are an irregular, nominal-measure (categorical) time series, events must be aggregated before the data can be used by standard statistical programs such as SPSS and SAS or graphical displays such as spreadsheets; these expect a regular, interval-measure (numerical) time series. This transformation is usually done by mapping each event code to an interval-level scale (for example, Goldstein 1992), and then aggregating the data by actor-pair and week, month or year using averages or totals.

---

[5] Null coding eliminates words that might be confused with codeable actors—for example JORDAN_RIVER—and verbs that should only be coded when they occur in combination with a phrase. Determining null codes is a substantial part of dictionary development; see Schrodt, Davis and Weddle (1994).

[6] KEDS provides a wide variety of formatting options for its output; typically the output file is read into a spreadsheet or database program for further processing.

It is *possible* to do the aggregation using the data transformation facilities of a statistical program. However, this process tends to be very slow and awkward, particularly when dealing with a large number of actor pairs. We have developed a flexible aggregation program, KEDS_Count, to automate this process; the program and its documentation are available on the KEDS web site. In contrast to the text reformatting programs, which need to be customized, KEDS_Count should handle most situations involving the aggregation of event data into a time-series.

"FREQUENTLY ASKED QUESTIONS"

Where can I get the program?

The most up-to-date version of the program—as well as the manual, utility programs, various coding dictionaries, data sets and so forth—is available from the KEDS web site:

http:// raven.cc.ukans.edu / ~keds

If you do not have access to the web, we can send you the program and other information on a disk.

Are you planning to sell a commercial version?

No.

What are the system requirements?

We have used KEDS in a variety of Macintosh configurations, including an SE, SE/30, original II, II with a DayStar Turbo 30 accelerator, IIsi, IIsi with a DayStar Turbo 40 accelerator, Quadra 900, LC, Powerbook 160, Powerbook 520c, and Power Macintosh 7100; we've also used it under Systems 6.0.5, 6.0.7, 7.1, and 7.5. The suggested application memory size is set at 2048K so KEDS should run on Macs with 4Mb or more of memory under System 6, or 8Mb under System 7.[7]

How fast is the coding?

Roughly 15 events per second on a PowerPC 7100/80 or on a Mac IIsi running with a DayStar 68040 50Mhz accelerator. Coding our 80,000 event 1979-1996 Middle East data set using a 4000-phrase *.Verbs* dictionary and 700-phrase *.Actors* dictionary requires about two hours.

Will the program run on Windows95?

The KEDS project is part of the valiant but desperate rear-guard effort to prevent William Gates III from controlling the entire known universe. Providing KEDS exclusively on the Macintosh platform is our small contribution to that cause.

Actually, we intend to port the program to Windows95 at the point where we believe that all of the major bugs have been located in the Macintosh version. This will occur real soon now.

How long does it take to develop dictionaries?

This depends entirely on how much your event coding scheme differs from the scheme used in an existing dictionaries. We estimate that around two person-years (4000 hours) went into the development of our Middle East dictionaries, and a comparable level of effort has gone into the PANDA dictionaries. However, this involved a lot of dead-ends and the effort was integrated into

[7] The 2Mb requirement is actually quite conservative in order to allow for memory-intensive operations such as indexing; the program will probably run reliably in about 1.4Mb of memory.

debugging the program itself (and in the case of PANDA, refining the coding scheme). Translating the WEIS codes into the BCOW system, in contrast, took one person less than a month. The existing dictionaries probably contain most of the English-language vocabulary relevant to coding political events, but usually one needs to work with actual news reports to determine the best association of phrases and codes.

Other work in automated text processing of reports of political events (ARPA 1993; Linert & Sondheim 1991) indicates that dictionaries on the order of about 5,000 words are necessary for relatively complete discrimination between political events described by news media sources in the English language. The KEDS and PANDA dictionaries are somewhat smaller than this—about 4000 phrases—and a dictionary focusing on only a small subset of behavior might be substantially smaller.

Can KEDS code full stories in addition to lead sentences?

In the case of Reuters, yes—Phillip Huxtable (see Schrodt, Huxtable & Gerner 1996) has been using full-story coding to generate a data set for West Africa. KEDS can pass a pronoun reference across sentences, and in Reuters stories a pronoun at the beginning of a sentence usually refers to the first actor in the previous sentence. Full stories have much more redundant and irrelevant information than do lead sentences, so it is necessary to use different pre-processing and post-processing filters for full-story coding than are used in lead sentences.

Based on our validity tests, lead sentences alone provided sufficient coverage in the Middle East, an area closely monitored by the international media. As noted earlier, our data set based on Reuters leads provided almost three times the density of events found in human coding of full stories from the *New York Times*, in part because Reuters tends to break up any major story on the Middle East into several small stories. Huxtable found the opposite pattern to hold in West Africa—Reuters tended to combine multiple, virtually unrelated events into a single story—so full-story coding provided a considerable increase in density.

Is KEDS useful for projects other than coding event data?

Maybe, and maybe not. KEDS is optimized for event data coding, but has a number of additional string-recognition features that can be applied to content analysis projects.[8] However, several important features found in most general purpose content analysis packages—notably Boolean searches and proximity-based searches—are not implemented in KEDS. Unless your coding scheme relies heavily on the subject-verb-object structure of a sentence or needs to handle compound nouns, you are probably better off with a content analysis program; Evans (1996) provides a review of these.

Will KEDS work for languages other than English?

In principle, yes, and in limited practice, yes. An early version of the program was used to code German-language reports quite successfully, but the specialized sections of the parser that handled German were not maintained in later versions. Depending on the language, it may be possible to use KEDS' rules facility to transform the structure of a sentence to something resembling English, but we have not tried this. The PANDA project is currently working on a machine-coding system that works in Spanish.

---

[8] These features, known as "ISSUES" in KEDS, were designed to code information about the context of an event—for example distinguishing food riots from ethnic riots—and the extract specific numerical information from the text, such as the number of people arrested in a demonstration.

Is the program in the public domain?

Under the Bayh-Dole Act that governs technology developed with National Science Foundation funding, the KEDS program is the intellectual property of the University of Kansas. You may use and make copies of the program for educational, government and non-profit use without charge; the program can be posted to bulletin boards and included in software collections provided that a copy of the manual is included. If you wish to license the program or its source code[9] for commercial applications, please contact the University of Kansas Technology Transfer Office (phone: 913-864-3302; fax: 913-864-5272).

Because you are undoubtedly wondering...

The term KEDS is an acronym for "Kansas Event Data System." The software is in no way connected to—nor could it possibly be confused with—a trademarked brand of footwear with a similar name.

APPLICATION: MEDIA FATIGUE IN COVERAGE OF THE PALESTINIAN INTIFADA

The issue of media bias has been a frequent concern in event data analysis.[10] For instance, an early study by Doran, Pendley, & Antunes (1973) found that coding regional sources produced dramatically different results than coding the *New York Times* when measuring the level of domestic conflict in Latin America. The issue we address here is somewhat different: the effects that occur over time in the coverage of a protracted conflict—the Palestinian *intifada*—by a single source.

It is quite plausible that coverage of a protracted conflict is unusually high when hostilities first break out and then declines as reporters, editors, and readers become bored with the issue. Alternatively, after the initial wave of reports, there may be a relatively fixed level of coverage for any given protracted conflict, regardless of the actual events on the ground. A second factor that may influence the extent of media coverage is competition from other new events. For example, after visually examining the event series dealing with the *intifada*, we speculated that there was a drop-off in coverage of the *intifada* by the *New York Times* in the summer and fall of 1989 when attention turned to the collapse of communism regimes in Eastern Europe (Schrodt & Gerner 1994).

Investigating media fatigue is inherently problematic because one must have an independent benchmark for what *actually* occurred, which in general cannot be ascertained in the absence of media coverage. (In any case, one cannot use media coverage to measure media fatigue.) In our analysis, we deal with this problem in two different ways. First, we use the independent measure of Palestinian deaths caused by Israeli military forces and their agents during the *intifada* (PHRIC 1993) and compare this with a net conflict/cooperation measure aggregated using the Goldstein scale, and with the number of force events coded from Reuters. Second, we compare the events generated from Reuters with the events that Tomlinson (1993) generated from the *New York Times*. We will compare these sources both directly and in terms of their correlations. This allows us to see whether Reuters (which often produces reports that do not end up in the newspapers) reflects events "on the ground" more closely than does the *New York Times*.

---

[9] The program is written in Think Pascal and contains about 16,000 lines of code.

[10] An earlier version of this research was presented in Gerner and Schrodt (1994).

As Table 2 shows, the overall correlations between the measures of conflict are high when the entire series is considered; all of the correlations are significant at the 0.05 level.[11] We would not expect the number of deaths to be identical to the number of force events (WEIS categories 221, 222, and 223) because the force category also includes some non-fatal violent encounters and an incident resulting in multiple deaths might be reported as a single event.

Table 2: Correlation of Variables Used to Study Media Fatigue

	Reuters Goldstein	NYT Goldstein	Reuters "Force" Events
PHRIC Deaths	0.643	0.761	0.630
Reuters Goldstein		0.803	0.930
NYT Goldstein			0.794

Figure 2 compares the monthly time series for the total number of deaths reported by PHRIC and the number of "force" events in the Reuters event data. With the exception of a few months — most notably the summer and fall of 1989—the two series covary fairly regularly. In some cases, there are conspicuous outlying points due to incidents that generated a large number of deaths but a disproportionately small number of reports. Two such spikes are the initial deaths from the Rishon LeZion massacre in May 1990 (and additional deaths that occurred with the suppression of demonstrations following the massacre) and the Haram ash-Sharif massacre and aftermath in October 1990.

In order to explore the issue of whether this pattern has changed over time, we regressed the number of deaths on the net conflict score produced by both our data set and Tomlinson's human-coded WEIS set based on the *New York Times*, and then looked at the residuals of this. If a residual is greater than zero, it means that the number of deaths is greater than the average being reported by Reuters and the *Times*; if the residual is negative, the number of deaths is lower.

These results are reported in Figure 3. As expected, the summer and fall of 1989 are conspicuous positive outliers, where deaths appear to have been under-reported. While we initially thought this problem affected the *New York Times* more than Reuters, it appears to influence both sources. Under-reporting also occurs in March and April 1988, although this is followed immediately by two months of over-reporting.

[11] The Goldstein scores have been multiplied by -1 so the correlations are positive; this is our "net conflict" variable.

Figure 2: Deaths versus "Force" Events by Month

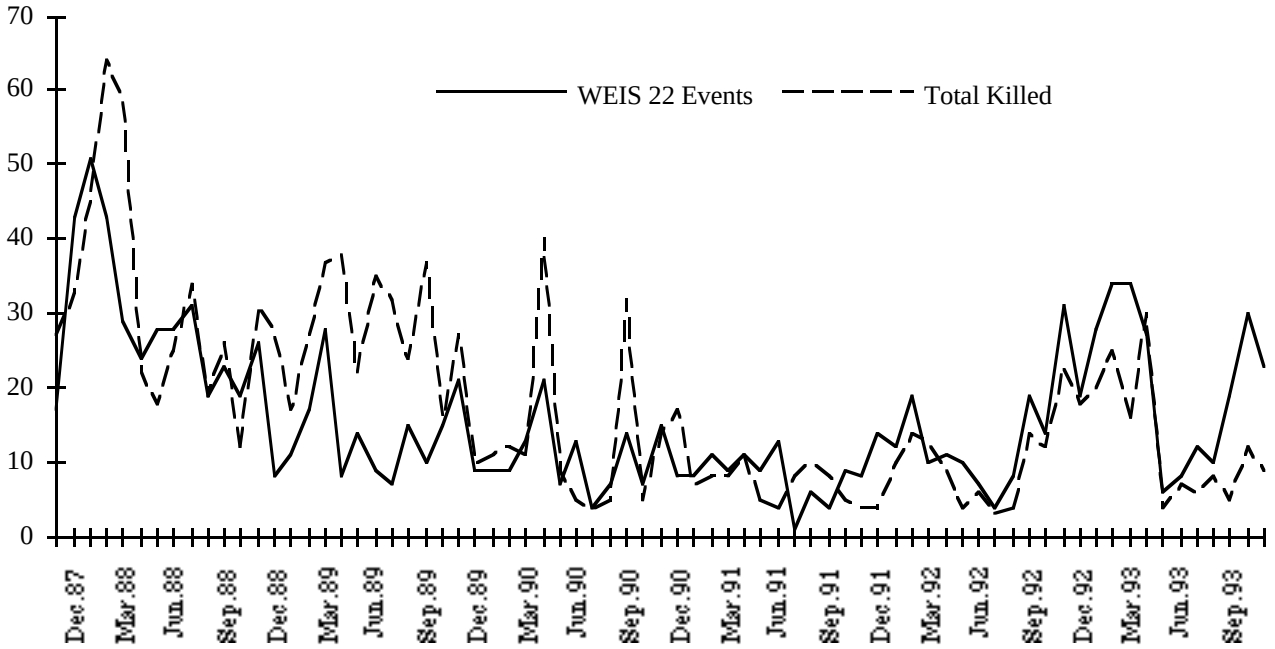
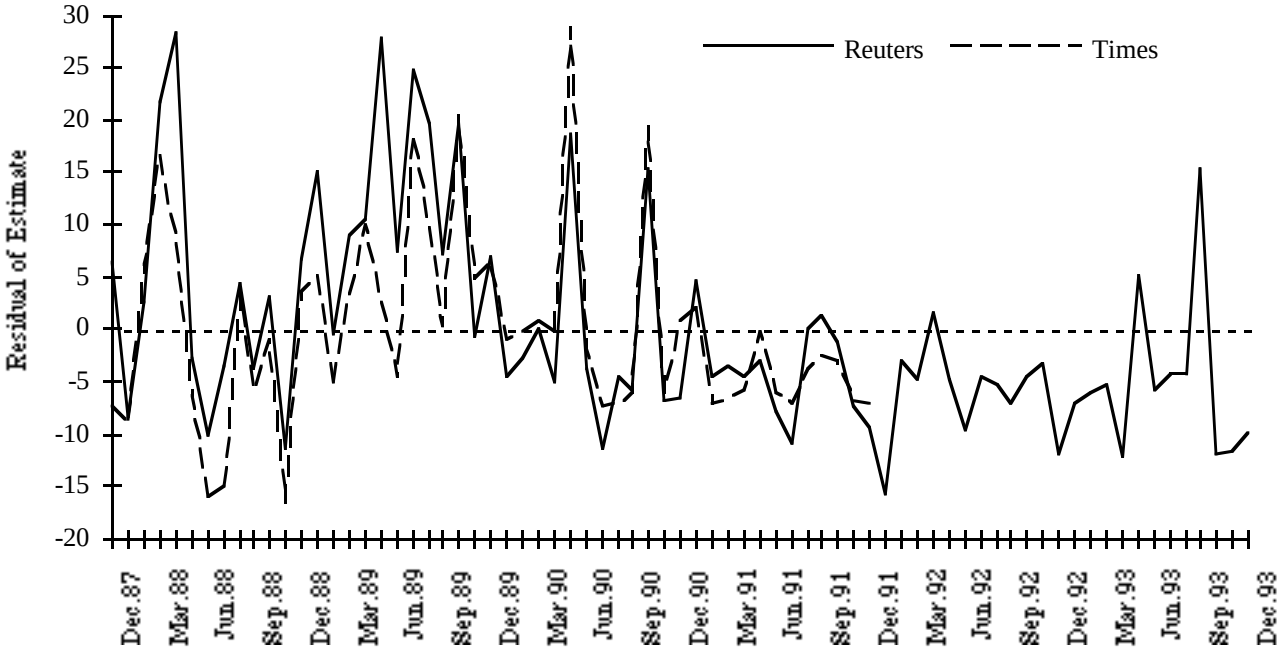
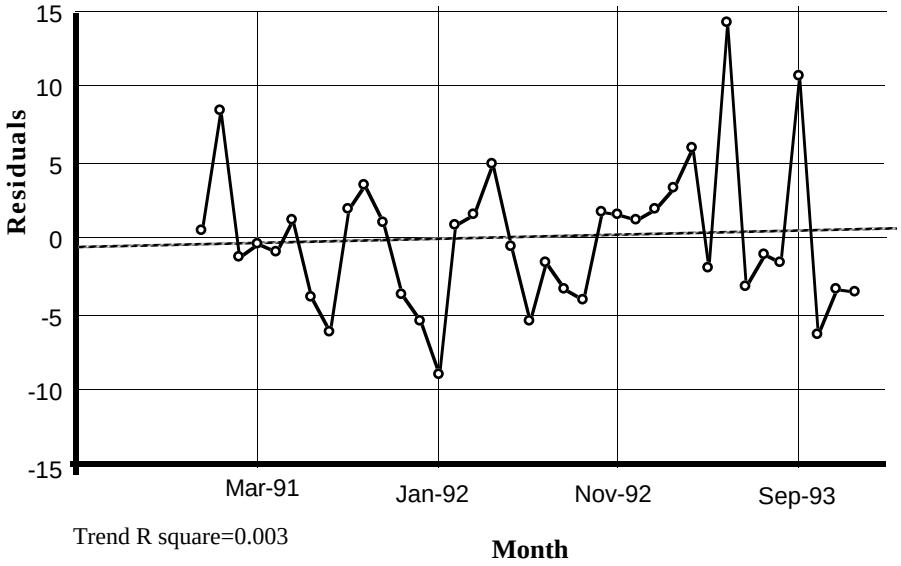


Figure 3: Deaths versus Net Conflict Score by Month



The apparent pattern of over-reporting (negative residuals) by Reuters in the second half of the series is an artifact of the outlying points in the 1989 and 1990. Figure 4 shows the residuals for only the period December 1990-December 1993; the dotted line is a trend determined by regression. This trend is almost flat, so we conclude that since December 1990, the relationship between deaths and the net conflict measured by Reuters reports has not been consistently increasing or decreasing.[12]

Figure 4:Residual Relationship of Deaths and Net Conflict, December 1990 - December 1993



Somewhat to our surprise, the coverage by the *New York Times* follows the trends of Reuters very closely; early 1988, summer and spring 1989 and the two 1990 massacres are the primary outliers. The *Times*-based data only goes to December 1991, but by then the *Times* seems to have settled down to about the same constant level of reporting as Reuters. In this sense the two sources are far more similar than we anticipated.

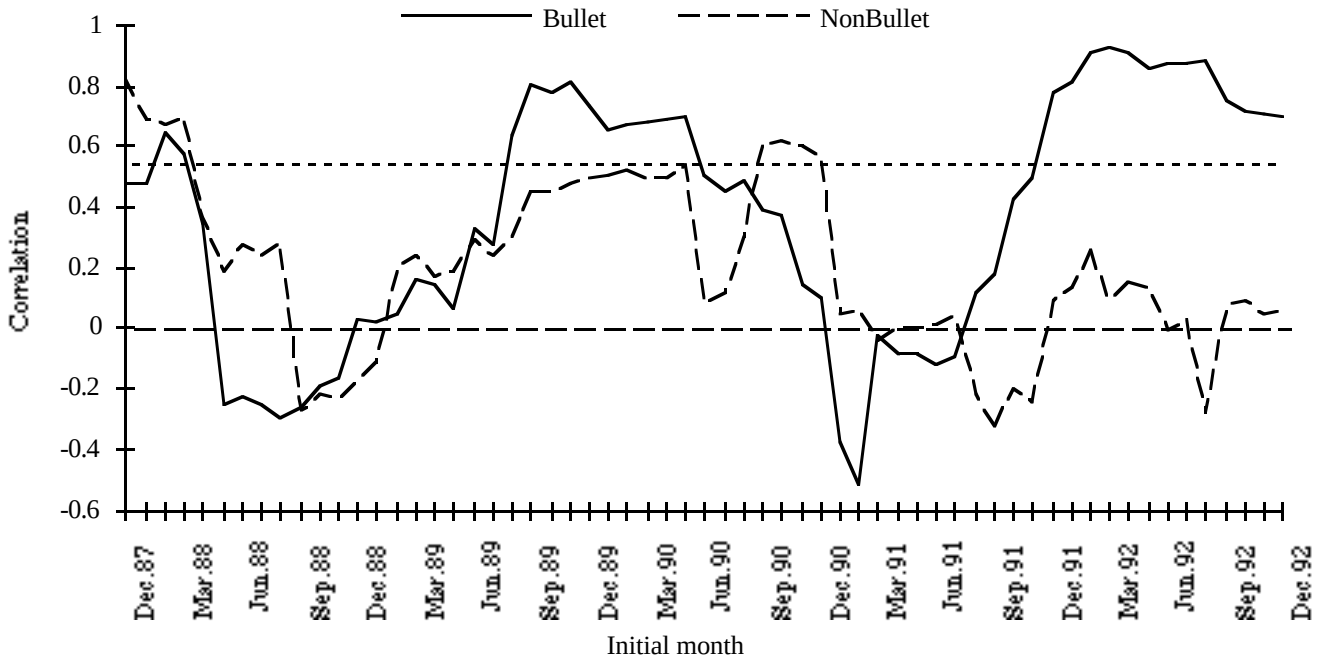
As Figure 4 shows, even when a trend is absent from the media coverage, there is still a substantial scatter in the data. While the overall correlation between the variables is significant, the pattern of correlation may change over time. In other words, the media may cover some periods more consistently than others. To explore this issue, we computed correlations between the various variables for moving 12-month intervals. The 12 month period was chosen because the Israeli-Palestinian data show annual seasonality (Schrodt & Gerner 1994). The results of these calculations are shown in Figures 5 through 8. In each of these figures the labels on the X-axis show the *initial* month of the 12 months of data used to compute the correlation. For example, the point labeled "Aug-89" is the correlation of points from August 1989 through July 1990. The dotted line at 0.57 is the approximate 0.05 significance level.

Figures 5 and 6 show the correlation between net conflict reported by Reuters and the *Times* and two measures of death reported by PHRIC, bullet and nonbullet (primarily deaths from beatings

[12] From eyeballing Figure 2, this constant level also probably applies to the spring of 1988 and most of 1990 except for the outliers in June and October.

and inhalation of tear gas).[13] The moving correlation shows a very different pattern than the comparisons of levels, particularly with respect to the effects of media attention on Eastern Europe. Both the *Times* and Reuters show a high correlation during the first 16 months of the *intifada*, but after this the 12-month correlation drops quite dramatically for about a year in the case of Reuters and for nine months in the case of the *Times*. Most of these periods of low correlation include some of the months in the second half of 1989, although the drop-off in correlation begins well before that time. Beginning about January 1989 for the *Times* and August 1989 for Reuters, the correlation increases and stays significant for about a year in both cases before dropping off again for almost two years. In the case of Reuters, this second decline corresponds closely to the Iraq-Kuwait crisis of 1990-1991. The drop-off occurs much earlier for the *Times*, and as we had speculated, the correlation is particularly low in the second half of 1989. The *Times* series is increasing as it ends; the Reuters series reaches a highly significant and fairly steady level after the middle of 1991. The correlations for non-bullet deaths are substantially lower than those for bullet deaths but the general patterns are similar, except for the post-1991 period in Reuters.[14]

Figure 5: Reuters 12 Month Correlation of Net Conflict and Deaths



[13] The moving correlations for the WEIS "Force" events are almost identical to those for net conflict; we did not have the force events tabulated for the *Times* data so we are using the net conflict scores in both cases to make them comparable.

[14] Non-bullet deaths also declined as a percentage of total deaths and relatively few occurred during the post-1991 period.

Figure 6: Times 12 Month Correlation of Net Conflict and Deaths

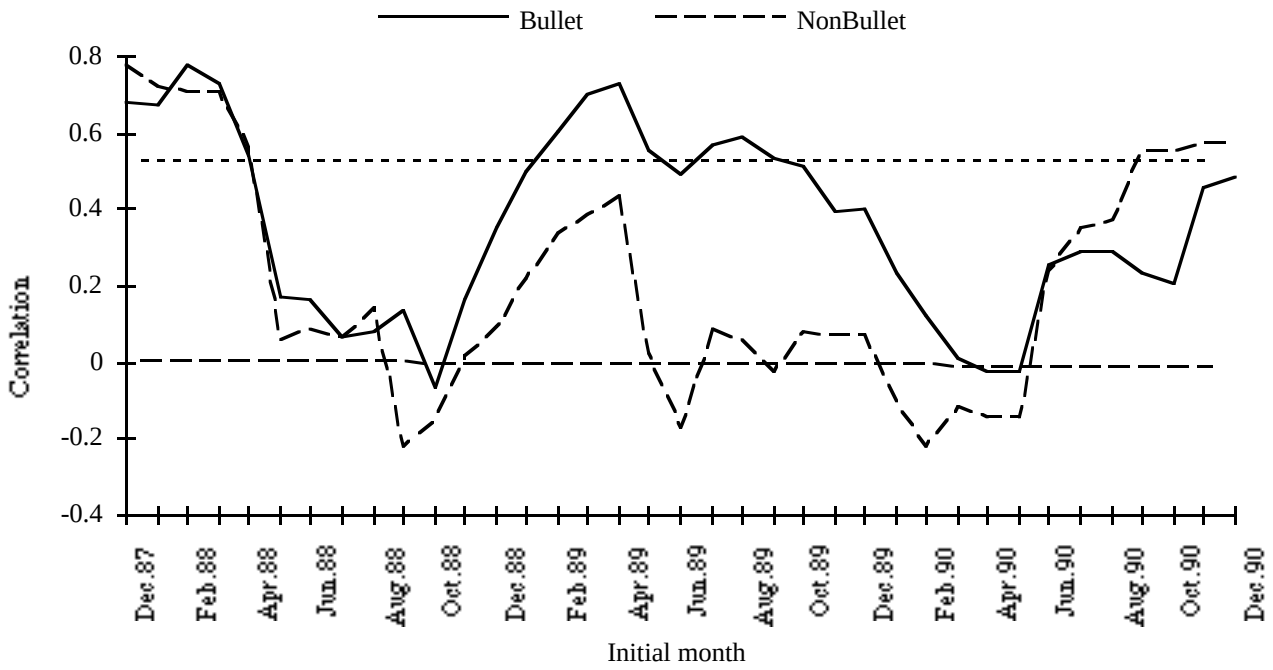


Figure 7: 12 Month Correlation of Reuters and Times, ISR->PAL

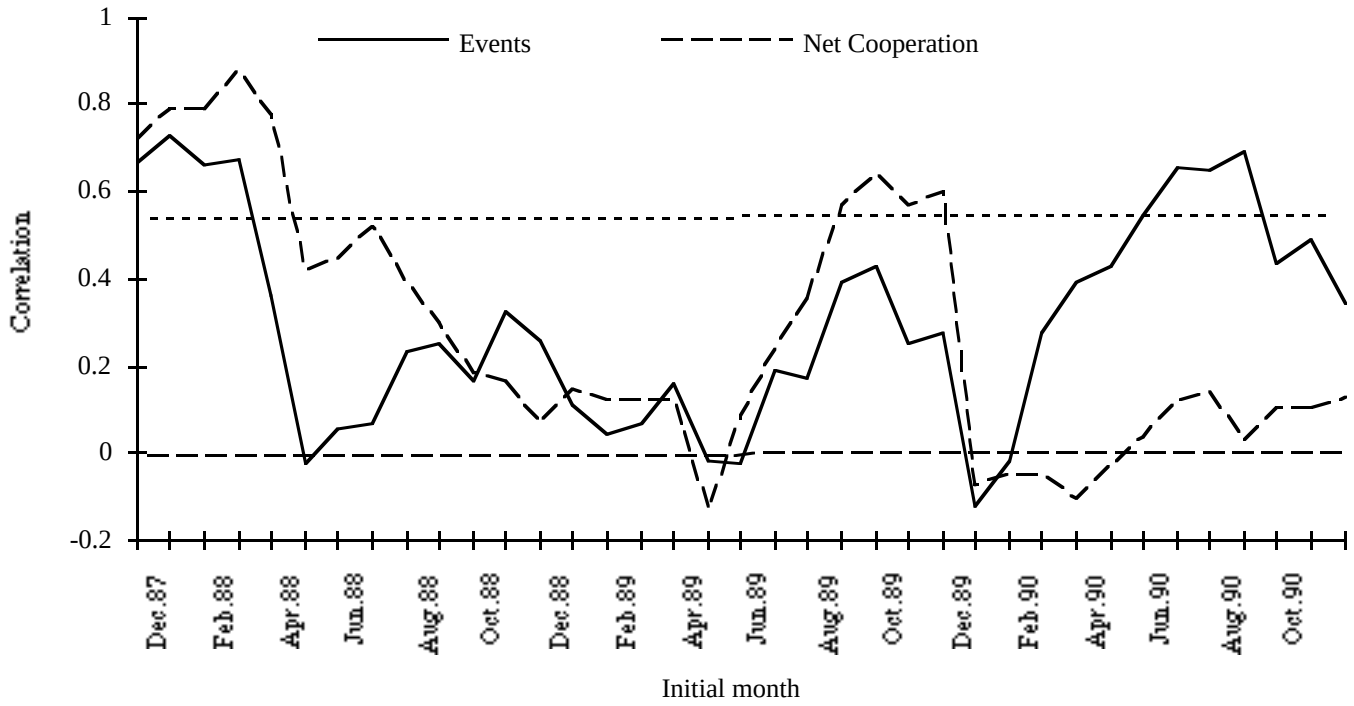
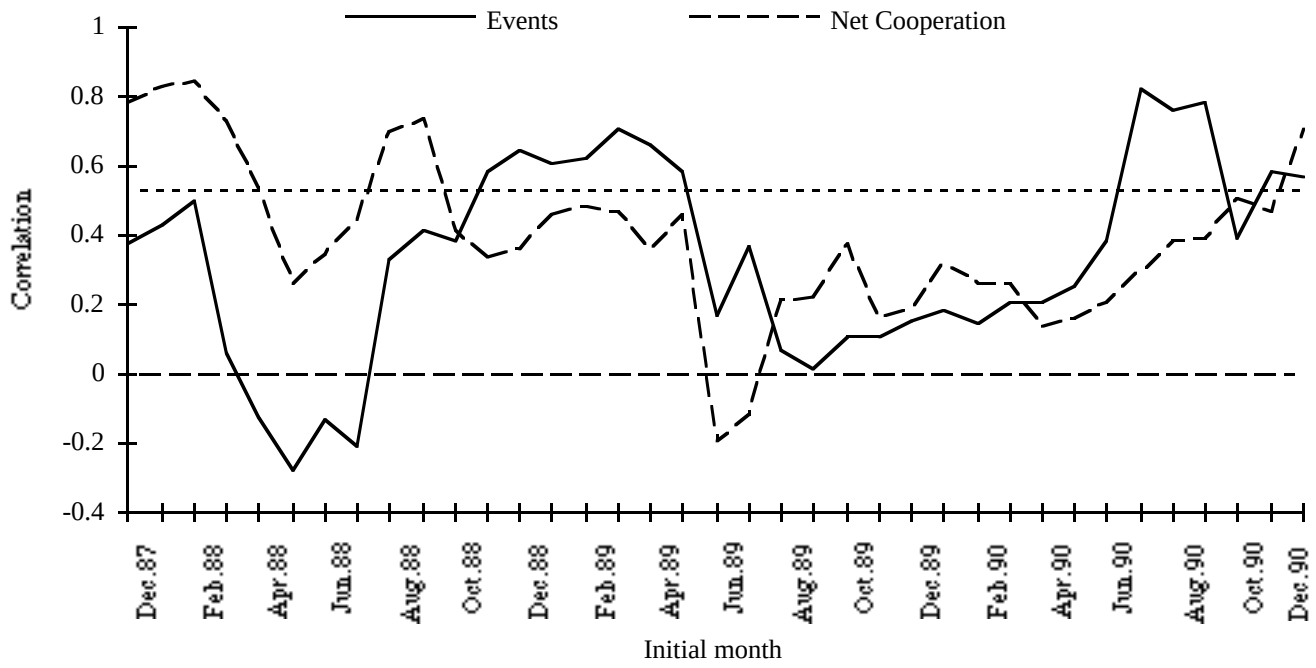


Figure 8: 12 Month Correlation of Times and Reuters, PAL->ISR



Figures 7 and 8 show the moving correlations between the number of events and net cooperation computed from Reuters and the *Times* for both the Israel->Palestinians and the Palestinians->Israel dyads. For the first two years of the reported correlations (covering the first three years of the data), the changes in the Israel->Palestinians pattern follow much the same pattern as the Reuters correlation with deaths. After January 1990, there is a rapid drop in the correlation in terms of the net cooperation measure, although the correlation with the number of events increases near the end of the period. These final months correspond to the Iraq-Kuwait crisis and the decline in the correlation may be due to Reuters and the *Times* having different reporting priorities.

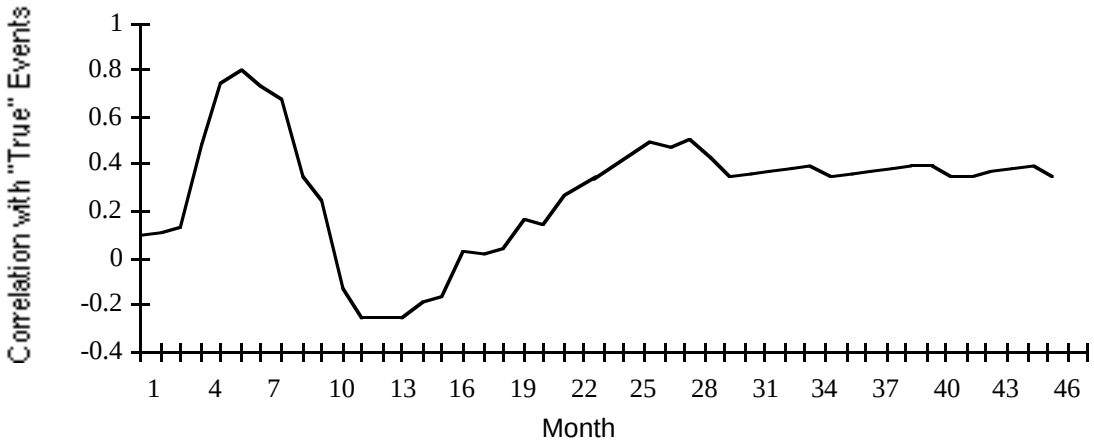
The pattern of correlation in the coverage of the Palestinians->Israel directed dyad is quite different from that of the Israel->Palestinians dyad. The Palestinians->Israel correlation across the entire data set is 0.63, as opposed to 0.80 for Israel->Palestinians, so the levels are generally lower. The correlation in the net cooperation measure declines much more slowly than in the Israel->Palestinians case, and does not show the increase in the months starting about August 1989. The slower decline in Palestinians->Israel may be due to what is known in journalism as the "man bites dog" effect: After several months of reporting Israeli uses of force against Palestinians (which were quite common), this phenomenon was no longer considered newsworthy, whereas Palestinian uses of force against Israelis—relatively unusual in the early stages of the *intifada*—continued to be newsworthy because of their novelty. The correlation at the end of the period is significant for both the number of events and net cooperation score, whereas there is almost no correlation in net cooperation in the Israel-> Palestinians directed dyad.

If the patterns of coverage of the *intifada* can be generalized to other protracted conflicts, we can draw three general conclusions about the issue of "media fatigue." First, all of the correlations show a short-term rise and then decline in the initial year of the *intifada*. This is the strongest pattern in the analysis and is consistent with the media becoming very interested in the conflict, covering it intensely for six months to a year, and then losing interest. As attention to the *intifada* declines, the correlation between the two sources we looked at also declines; in other words, as coverage becomes less thorough, it also becomes less consistent. For a short term of perhaps eighteen months, media fatigue seems to operate.

The long term, however, shows quite a different pattern and two other factors appear to be relevant. First, the influence of competing stories—the Iraq-Kuwait crisis and the collapse of communism in the Soviet Union and Eastern Europe—seem more important than a general fatigue pattern. These competing stories have a particularly strong impact on the correlation between the two sources: Once again, as interest declines, so does consistency. In the absence of a competing story, however, the fatigue effect appears to end and coverage reaches a steady-state, at least in the final three years of the Reuters data. There are insufficient data to ascertain whether the same steady-state would be found in the *Times* data.[15]

In short, the media fatigue effect in protracted conflicts may look something like Figure 9: an initial interest followed by a slump followed by sustained interest. The timing of the phases may be differ somewhat between sources: for example, the *Times* seems to have sustained its initial interest in the *intifada* longer than did Reuters. Because of the impact of the Soviet Union-Eastern Europe and Iraq-Kuwait, the pattern we found in the *intifada* was not as clean as the hypothetical pattern.

Figure 9: Hypothetical Media Fatigue Pattern



The media fatigue results are reassuring in some respects, but they also point to cautions for future research. Neither the short-term pattern of interest-then-decline nor the impact of competing news stories was unexpected, although in our earlier work we underestimated the impact of the second factor, particularly on Reuters. In the *intifada* this initial period of fluctuating attention is followed by a long period of sustained and quite consistent news coverage, which argues in favor of the use of journalistic sources in monitoring protracted conflicts. With sufficiently sophisticated models and the use of multiple sources, it should be possible to compensate (at least partially) for coverage effects. For instance, one could use the reduced correlations in reports by multiple sources (e.g. Reuters, Agence France Presse, and the *New York Times*) to signal periods of reduced coverage.

---

[15] There is a significant decline in *intifada*-related deaths through 1990, 1991, and the beginning of 1992 (with the exceptions already mentioned). Thus, it is difficult to determine whether the steady-state level of media coverage had a high correlation to the level of deaths because of the absence of media fatigue or because there was less to report. We lean toward the former explanation because Reuters *did* pick up on the sharp increase in deaths during the latter months of 1992 and the first half of 1993. Unfortunately, we do not have reliable PHRIC data beyond 1994 and it is not possible to investigate this further.

CONCLUSION

The generation of event data by machine, at least within the WEIS framework, is not only possible but relatively easy. While the methods used with Reuters would not work on less structured text such as political rhetoric, coding Reuters alone provides a substantial improvement over the existing event data sets.[16] Based on our experience with KEDS, relatively simple programs running on inexpensive personal computers can be used to do a significant amount of processing of this type of data. We view KEDS as a first step, not the last word, and we would encourage others to work on these problems, as the potential rewards to social science research appear quite substantial.

Machine coding may be useful for generating specialized data sets from regionally-specific or historical sources that might not otherwise be coded due to the labor involved. The preparation of the input text can be done by individuals without special training, such as work-study students or university word-processing centers; specialists can focus on the development of the coding dictionaries. If machine-coding systems can be in languages other than English—and there is every reason to assume that this is possible—then the expertise of multi-lingual coders can be concentrated dictionary development.

Over the past several years, a substantial number of sources of news text have become available on CD-ROM. For example, the full text of *The New York Times* is available on CD-ROM for 1991 and all subsequent years; the Dialog information service also provides the full text for a number of newspapers on CD-ROM. University Microfilms Inc. produces a CD-ROM that contains abstracts of a number of US newspapers, including *The New York Times*, *Wall Street Journal* and *Washington Post*. Since the marginal cost of producing a CD-ROM from electronically-stored text is very low, presumably Reuters and other news agencies will provide this service in the near future.

While substantially faster than human coding, KEDS is still not quite fast enough to allow interactive experimentation with event coding schemes. Ultimately one would like to have a system where a researcher could change event definitions see these coding changes reflected, within a minute or two, in a statistical summary of the data, much as one can currently experiment with numerical transformations and subsets in SPSS or SAS. With source text files edited to eliminate uncodeable records, faster microprocessors, inexpensive parallel processing using networks and coprocessors, and greater efficiency within the KEDS program itself, this objective is probably attainable within the next five years.

---

[16] It is possible to create systems to process more complex texts, but only if one limits their substantive domain. For example Alvarado's OpEd program (1990) "understands" editorials on issues in political economy; Kolodner's CYRUS (1984) deals with the official activities of Secretary of State Cyrus Vance and Pazzani's OCCAM (1988) deals with international economic sanctions. The performance of such systems is quite impressive but they are dependent on a large amount of domain-specific knowledge and can only work with text dealing directly in that domain. OpEd, for example, required prior editing "to remove . . . issues that fall outside of OpEd's process model, such as understanding references to historical events, completing analogies and handling sarcastic or humorous statements" (Alvarado, 1990:3).

KEDS, in contrast, deals with unedited source texts covering a very wide range of political behavior. KEDS cannot deal with that material in much depth, but its processing appears adequate for event data coding. The tradeoff between complexity and specificity is generally encountered in technology: one can either perform complex tasks in a limited domain or simple tasks in broad domains, but not both. Almost any hammer can be used to pound almost any nail, but you can't use the remote control for a VCR to program a microwave oven.

SUGGESTED READINGS

This annotated bibliography gives citations to the primary published papers from the KEDS project, as well as some surveys of contemporary event data analysis and computational methods for processing natural language.

KEDS

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

- Description of the DDIR-sponsored KEDS research; includes tests on German-language sources and a foreign affairs chronology.

Schrodt, Philip A. and Deborah J. Gerner. 1994. "Validity Assessment of a Machine-Coded Event Data Set for the Middle East, 1982-1992." *American Journal of Political Science* 38:825-854.

- Statistically compares KEDS data to a human-coded data set covering the same time period and actors.

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.

- A technical description of KEDS with an extended discussion of the types of problems encountered when machine-coding Reuters reports.

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

- Analysis of source bias problems comparing Reuters, *Agence France Press* and United Press International.

Event Data

Schrodt, Philip A. 1994. "Event Data in Foreign Policy Analysis" in Laura Neack, Jeanne A.K. Hey, and Patrick J. Haney. *Foreign Policy Analysis: Continuity and Change.* New York: Prentice-Hall, pp. 145-166.

- Textbook-level introduction to the general topic of event data analysis.

Duffy, Gavin, ed. 1994. *International Interactions* 20,1-2

- Special double-issue on event data analysis.

Merritt, Richard L., Robert G. Muncaster, and Dina A. Zinnes, eds. 1994. *Management of International Events: DDIR Phase II*. Ann Arbor: University of Michigan Press.

- Reports from the DDIR projects.

Computational Methods for Interpreting Text

Advanced Research Projects Agency (ARPA). 1993. *Proceedings of the Fifth Message Understanding Conference (MUC-5)*. Los Altos,CA: Morgan Kaufmann.

- Reports from a large-scale ARPA project on developing computer programs to interpret news reports on terrorism in Latin America; these use a variety of different techniques.

Evans, William. 1996. "Computer-Supported Content Analysis: Trends, Tools and Techniques." *Social Science Computer Review* 14,3: 269-279.

- Current survey of content-analysis programs

Pinker, Steven. 1994. *The Language Instinct.* New York: W. Morrow and Co.

- Excellent non-technical introduction to contemporary linguistics; extensive discussion of the problems of parsing English

Salton, Gerald. 1989. *Automatic Text Processing.* Reading, Mass: Addison-Wesley.

- General introduction to the use of computers to process text; covers a wide variety of methods.

BIBLIOGRAPHY

Advanced Research Projects Agency (ARPA). 1993. *Proceedings of the Fifth Message Understanding Conference (MUC-5)*. Los Altos,CA: Morgan Kaufmann.

Alvarado, S. J. 1990. Understanding Editorial Text: A Computer Model of Argument Comprehension. Boston: Kluwer Academic Publishers.

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, Bennett, Brad & Vogele, William 1994. Data development and interaction events analysis using KEDS/PANDA: an interim report. International Studies Association, Washington.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor and Kurt Schock. 1996. "Contours of Political Contention: Issues and Prospects for the Automated Development of Event Data." International Studies Association, San Diego.

Doran, Charles F., Robert E. Pendley, & George E. Antunes. 1973. A Test of Cross-National Event Reliability. *International Studies Quarterly* 17:175-203.

Evans, William. 1996. "Computer-Supported Content Analysis: Trends, Tools and Techniques." *Social Science Computer Review* 14,3:269-279.

Gerner, Deborah J. and Philip A. Schrodt. 1994. "Foreign Policy Interactions in the Middle East: An Initial Examination of Three Cases of Conflict." International Studies Association, Washington.

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

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

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

Kolodner, J. L. 1984. *Retrieval and Organizational Strategies in Conceptual Memory: A Computer Model*. Hillsdale, NJ: Lawrence Erlbaum.

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

Lehnert, Wendy and Beth Sundheim. 1991. "A Performance Evaluation of Text Analysis." *AI Magazine* 12:81-94.

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

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

Palestine Human Rights Information Center. 1993. *Human Rights Violations Under Israeli Rule During the Uprising*. Washington, DC: PHRIC.

Pazzani, M. 1988. *Learning Causal Relationships: An Integration of Empirical and Explanation-Based Learning Methods*. Ph.D. dissertation, University of California, Los Angeles.

Schrodt, Philip A. and Deborah J. Gerner. 1993. "Statistical Patterns in a Dense Event Data Set for the Middle East, 1979-1992." Midwest Political Science Association, Chicago.

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

Schrodt, Philip A. and Deborah J. Gerner. 1996. "Using Cluster Analysis to Derive Early Warning Indicators for the Middle East, 1979-1996" American Political Science Association, San Francisco.

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.

Schrodt, Philip A., Philip A. Huxtable & Deborah J. Gerner. 1996. "Events Data and the Analysis of Political Behavior: The Middle East and West Africa, 1979-1995." International Studies Association, San Diego.

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