Working with Event Data: A Guide to Aggregation Choices[1]

James E. Yonamine

Department of Political Science
Pennsylvania State university
jxy190@psu.edu

"You live and die by your coding rules."
-David Singer

Event data is becoming increasingly prominent in the peer-reviewed conflict literature, especially among studies that emphasize prediction. With the rapid increase in freely available electronic text, this trend is likely to continue. Currently, most major event data datasets, including the Integrated Conflict Early Warning Systems (ICEWS) project, the 10 Million International Dyadic Events dataset, (http://gking.harvard.edu/publications/10-million-international-dyadic-events) and all of the Penn State Event Data Project datasets (http://eventdata.psu.edu/data.html) provide information in a [who-did what-to whom, when] format at the daily level, that resembles the data presented in the table below:[2]

Table 1: Example of Raw Event Data Content

Date	Source	Target	Action code
50305	INSEP	INDGOV	831
50510	INDREB	INDGOV	40
50823	IND	PAK	141

Unfortunately, event data in the raw is unsuitable for both theory building and for empirical models. Thus, researchers must convert raw event data to a more usable format. The fine grained structure of raw event data provides researchers with a tremendous amount of flexibility regarding aggregation techniques across the three primary dimensions of manipulation: actors, actions, and time. Counting only the aggregation strategies used across these three dimensions in the extant literature, over 500 potential theoretically justifiable combinations exist for every research design.[3] Despite this, studies utilizing event data rarely provide rigorous discussion of aggregation techniques across all three dimensions.[4] This is especially problematic, as Shellman (2004*a*); Alt,

[1]Will Lowe has created an excellent event data aggregation package for r called "events". It is available on his website: http://www.williamlowe.net/software/. This package is able to perform most of the aggregation techniques discussed in this paper. For those who prefer STATA, I can provide scripts via e-mail upon request.

[2]The "Action Code" uses the Conflict and Mediation Event Observation (CAMEO) typology.

[3]This is a low-ball estimate and only assumes 10 types of actor aggregation when many multiples of that exist.

[4]Some exceptions include Shellman and Stewart (2007); Moore (1998); Gleditsch and Beardsley (2004); D'Orazio, Yonamine and Schrodt (2011).

King and Signorino (2001); Freeman (1989); Thomas (2002) have all demonstrated that empirical findings in event data studies often change when analyzed at different levels of temporal aggregation. Given the large number of potential aggregation choices and their potential to effect empirical findings, it is critical that event data studies in the future should clearly articulate their aggregation strategies across the three key dimensions. To facilitate how scholars make and document aggregation choices in the future, I provide an overview of the existing aggregation techniques, highlighting strengths, weakness, and best practices when applicable.[5] I argue that every study using event data should clearly discuss their actor, action, and temporal aggregation techniques, and the remainder of this paper follows that order.

1 Actor Aggregation

Raw event data provide information regarding the actors involved with the action, generally in terms of a 'source' and a 'target,' although some actions are non-directional. Among the three major dimensions of raw event data aggregations (actors, actions, and dates), actor identification techniques have the highest degree of variation between different data sets and taxonomies. Some coding schemes, including the original COPDAB and WEIS, use a 3-character identifier for all unique actor groups while others utilize broad numerical codes. In this section, I focus on the CAMEO taxonomy (also used by the ICEWS project), which uses hierarchically constructed actor identification codes. Under this scheme, each actors receives a 3- to 12-character code. At a minimum, every actor receives the 3-character country code, but additional 3-character identification codes are appended in a hierarchical fashion based on the level of description provided in the text. Thus, actor codes in CAMEO range from the general [IND] (i.e. India) to the specific [INDGOVOPPPTY] (i.e. Indian Government Opposition Party). Given this, researchers must determine the actors of interest between whom an event must occur in order for that event to be included in their study's empirical models. Additionally, of the three areas of aggregation, scholars tend to provide the least information about their choices about actors. Though I am unaware of any study that replicates existing event data analyses using different actor aggregation techniques, it is likely that doing so could have dramatic effects of empirical findings.

At a minimum, it is critical that scholars focus on events involving at least one actor affiliated with a country of interest. Substantively, justification for this minimal level of actor aggregation is clear; a study focusing on Israeli-Palestinian conflicts would not want to include events between Aceh rebels and the Indonesian army, as these are not relevant to the conflict of interest. Although excluding Indonesian rebel activity is obvious, more

[5] This articles covered in this paper are by no means the total population of event data studies. However, I did my best to include all published papers in the last 10 years that use machine-coded event data to forecast political outcomes. If you know of an article I left out, please do not hesitate to contact me.

difficult decisions exist for this example, such as whether or not to include events between members of the Lebanese and Syrian armies or between the governments of the United States and Iran. For example, studies interest in inter-state relations should make explicit whether all actors form a certain country are included (i.e. rebels, civilians, police) or only specific, policy relevant position (l.e. legislators, president, party members).

In addition to limiting actors by country, scholars interested in intrastate dynamics should be explicit about the the treatment of actor codes beyond the 3-character country identification. It is not sufficient for a study to merely state that it is focusing on "sub-state actors" or "relevant domestic actors". Instead, it is crucial that scholars clearly state how they treat specific secondary, tertiary, or even quaternary 3-character codes. For example, a study interested in domestic rebel groups should clearly illustrate the specific actor identification codes that they treat assume to indicate a rebel. In some event data datasets, only one, catch-all code exists for all actors associated with a mobilized and armed opposition force, but for others, multiple 3-character codes exist.

Due to the lack of precedence of how to make and discuss actor aggregation choices, I provide a "best practice" discussion on actor aggregation from D'Orazio, Yonamine and Schrodt (2011) in the box below.[6]

> "Every coded event in the ICEWS dataset contains two actors: a source and a target. Because ICEWS uses the CAMEO coding ontology, each actor is coded using a three-tiered scheme. The first tier is provided for all actors and reflects national identity (for example, [CHN] for an actor identified as Chinese), which we require to be identical for both the source and the target. This ensures that we only analyze events occurring domestically. Additionally, the ICEWS dataset often includes a second tier (or 3-letter code) of information includes many sub-national level descriptions and, where applicable, a third-tier sub-sub-national descriptions. We drop the third tier and select only a relevant selection of actors from the second tier. Specifically, we build three main "classes" of actors based on their second tier categories for the events that occur domestically:
>
> - *Government*, which includes actors identified by ICEWS as:
>   - [MIL] – Military
>   - [POL] – Police
>   - [BUR] – Bureaucrats
>   - [POL] – Politicians
> - *Rebels*, which includes actors identified by ICEWS as:
>   - [INS] – Insurgents
>   - [SEP] – Separatists
>   - [REB] – Rebels
> - *Other*, which includes actors identified by ICEWS as:
>   - [CIV] – Civilians unaffiliated with another group
>   - [BUS] – Individuals identified as a business person
>   - [EDU] – Students and teachers
>
> The domestic events occurring between the government and rebel groups (whether labeled as insurgents, separatists, terrorists, or other categories) and between rebel groups and other non-governmental actors are the primary interactions comprising escalatory processes in intra-state conflict. Therefore, these are actors used. They comprise two classes of undirected actor dyads of interest: GOV-REB and REB-OTH. We omit all events not occurring between GOV-REB or GOV-OTHER actors with the same national identify."

[6] Gleditsch and Beardsley (2004) and Moore (1998) also provide comprehensive discussions of actor aggregation choices.

Additionally, I provide a table of the level of actor aggregation and spatial coverage for a number of select event data studies in order to facilitate literature reviews for future event data studies.

Table 2: Summary of level of actor aggregation in select studies

Article	Level of Actor Aggregation	Country/Region
Fordham (2005)	State	US, North Korea, Vietnam Syria, Iraq, Egypt
Pevehouse (2004)	State	Politically relevant dyads
Schrodt and Gerner (1997)	State	Lebanon Conflict
Sky (2000)	State and Sub-state	US and El Salvador
Goldstein (1991)	State	US and Soviet Union
Goldstein et al. (2001)	State	Middle Eastern states
Gleditsch and Beardsley (2004)	Sub-state	El Salvador, Guatemala Nicaragua
Schneider and Troeger (2006)	State	Israel-Palestine, Iraq war Gulf war, Ex-Yugoslavia
Brandt and Freeman (2004)	State	Israel-Palestine
Brandt, Colaresi and Freeman (2008)	State	Israel, Palestine, US
Shellman, Hatfield and Mills (2010)	Sub-state	Indonesia, East Timor Cambodia
Shellman (2004*a*)	Sub-state	Colombia and Afghanistan
Hämmerli, Gattiker and Weyermann (2006)	Sub-state	Afghanistan
Goldstein (1997)	State and IO's	Serbia, Bosnia, UN, NATO
Moore (1998)	Sub-state	Peru and Sri Lanka
Moore (2000)	Sub-state	Peru and Sri Lanka
Bond et al. (1997)	State and Sub-state	Poland, South Korea, China Yugoslavia
Brandt, Freeman and Schrodt (2011)	State	Israel and Palestine
Shellman and Stewart (2007)	Sub-state	Haiti
Shellman (2007)	Sub-state	Afghanistan
Shellman (2004*b*)	Sub-state	Chile and Venezuela
Shearer (2006)	State	Israel and Palestine
Schrodt and Gerner (2000)	State	Levant states
Kovar et al. (2000)	State	US and Iraq and Levant states
Stoll and Subramanian (2006)	State	Levant states
Schrodt, and Yilmaz (2003)	State	Middle East, Balkans West Africa
Schrodt and Gerner (2001)	State	Middle East and Yugoslavia
D'Orazio, Yonamine and Schrodt (2011)	Sub-state	Asia

2 Action Aggregation

In order to indicate the specific actions that occur between actors, event data datasets utilize an action typology, which provides structured numerical codes that correspond to a series out politically relevant events (see column 4 in Table 1). Within the peer-review published event data studies, CAMEO, WEIS, and IDEA are the most

commonly used action typologies. Since CAMEO and IDEA are built on the WEIS framework, these three typologies share similar characteristics. CAMEO, for example, utilizes 20 "cue categories", or classes of events, which contain different sub- and sub-sub categories. The list below CAMEO's hierarchical numerical code structure for COERCE, which is the 17^{th} "cue category". The two-digit, three-digit, and four-digit codes reflect the primary, secondary, and tertiary level of action code, respectively.

17: COERCE

- 170: Coerce, not specified below
- 171: Seize or damage property, not specified below
  - 1711: Confiscate property
  - 1712: Destroy property
- 172: Impose administrative sanctions, not specified below
  - 1721: Impose restrictions on political freedom
  - 1722: Ban political parties or politicians
  - 1723: Impose curfew
  - 1724: Impose state of emergency or martial law
- 173: Arrest, detain, or charge with legal action
- 174: Expel of deport individuals
- 175: Use tactics of violent repression

To the best of my knowledge, Kovar et al. (2000) is only one study has ever analyzed the complete, raw actions codes including the secondary and tertiary codes when provided. Aside from this, every study aggregates the action codes into some type of higher level variables. The vast majority of the extant event data literature converts the action codes for each event into coarse scale or count variables. Thus, this section is primarily dedicated to techniques used to build scale and count variables. However, it is important to note that a small number of studies using HMMs collapse secondary and tertiary action codes to the "cue category" level, then account for the number of each type of "cue category" event that occurs in a given temporal range (see Schrodt (1997), Bond et al. (2004), and Schrodt (2000)). Although this approach accounts for more information about the types of events that are occurring that any of the action aggregation techniques discussed below, it requires a methodological technique, like HMMs, particularly suited to handle a large number of sparsely populated independent variables.

2.1 Scale

Though rarely cited in modern event data publications, Azar and Sloan constructed the first event data scale (called the Azar-Sloan scale) in 1975 as part of the COPDAB project. This original scale was important because it placed events into a conflict-cooperation continuum and relied on a small sample of experts to determine degree of conflict/cooperation for each class of event — two ideas that are the foundation of the Goldstein Scale (discussed below), which dominates the current literature that uses scales. In brief, the Azar-Sloan scale consists of 15 cue categories, ranging from category 1, which reflects the most cooperational events to 15, which indicates the most conflictual. Azar and Sloan interviewed 18 international relations scholars to provided a weighted value for each cue category that reflects how may times more conflictual (for events 9-15) or cooperational (for events 1-7) each event is relative to category 8 events, deemed neutral. This resulted in categories 1-7 receiving values ranging from 6 to 92, with 92 indicating events that are 92 times more cooperational than the neutral category 8 events, and categories 9-15 receiving values ranging from 6 to 102.

In 1992, Goldstein used a similar operational approach (i.e. surveys of international relations professors) to build the conceptually similar Goldstein scale for the WEIS typology.[7] The Goldstein scale scores all WEIS actions on a -10 to +8.3 continuum, with -10 and +8.3 reflecting the most conflictual and cooperational events, respectively.[8] In order to determine the specific conflict-cooperation score to assign to each WEIS action, Goldstein surveyed either international relations faculty at the University of Southern California. Thus, despite the ubiquity of the Goldstein scale in the extant event data literature, the approach used to build the scale are questionable. Not satisfied with the informal survey approach that Goldstein originally used to build the Goldstein scale, Bond et al. (2003) utilized a web-based survey in order to rank all terminal nodes within the IDEA action typology on a conflict-cooperation scale. The result was conceptually similar to the Goldstein scale, but action scores ranged from -13 (for the most conflictual) to +7 (for the most cooperational).[9]

Despite the popularity of the Goldstein scale, scholars have recently built two additional scales based on coding typologies other than WEIS and WEIS compatible typologies like CAMEO and IDEA. First, the Intranational Political Interactions project is an alternative scaling approach that codes events into according to a unique typology of 10 classes of conflictual event and 10 classes of cooperational events. Each class of events takes on a score of +10 to +100 (in multiples of 10), which reflect the degree of cooperation conflict (see Shellman (2004*b*,*a*) for more information about the IPI scale). Second, the Violent Intranational Conflict Data Project (VICDP) collapses events into 15 categories, which escalate from most cooperational to most conflictual (see

[7] Prior scales for the WEIS typology proceeded the Goldstein scale but did not gain prominence in the literature. Goldstein (1992) provides a condensed summary of early efforts at building scales and counts from WEIS. Additionally, see Goldstein (1992) and Vincent (1983) for more thorough discussions of prior scaling approaches.

[8] Due to the similarities between WEIS and other taxonomies, the Goldstein scale is easily implemented on CAMEO and IDEA coded event data. See URL: http://eventdata.psu.edu/cameo.dir/CAMEO.scale.html for the Goldstein scale applied to CAMEO codes

[9] See Hämmerli, Gattiker and Weyermann (2006) and Bond et al. (2003) for additional information.

Moore and Lindström (1996) for additional information on the VICDP project). For example, category 1 is called "Agreement-Resolution" defined as the termination of the internal war, and category 15 is "Extensive war acts causing deaths, dislocation, and high strategic costs", which requires no further explanation.

Scholars who choose to use a scale must still make another important decision regarding how to aggregate the actual scores.[10] Common techniques are to calculate either the sum or the mean of the scaled scores, though both of these technique have potential shortcomings that I call the **mean problem**, **sum problem**, and the **single scale problem**.

- **Mean problem**:
  - A month with three -10 events occurring every day would have the same mean score as a month with only 1 ""-10 event per day. Since it is obvious that a month with 90 "-10 events" is more conflictual than a month with only 30 "-10" events, taking the mean score can lack external validity.
- **Sum problem**:
  - According to the Goldstein scale for WEIS data, "Issue order or command, insist, demand compliance" and "military attack; clash; assault" receives a -4.9 and -10 on the Goldstein scale, respectively. Consider two months, one with little dialogue but actual violence, and one with no violence but considerable negative dialogue. By summing the Goldstein values, the latter month could appear more conflictual than the first, even though it experience no actual conflict.
- **Single scale problem**:
  - Consider a day during which "noninjury descructive action" (a -8.3 on the Goldstein scale) and a "extend military assistance" (a +8.3 on the Goldstein scale) occur between the same actors. The sum and the mean of these two events equals 0, which is the same score that a day with no events receive. Theoretically, it is apparent that the nature of events occurring on a day comprised of purely neutral events and a day with a -8.3 event and a +8.3 event are fundamentally different than a day on which no relevant actions occurred. Additionally, many conflicts experience negotiation events (assigned small, positive scale values) during fighting (assigned large, negative scale values). However, these types of conflicts are impossible to differentiate from other conflicts with slightly lower levels of violence but no attempts at mediation or negotiation.

To overcome the single scale problem, some scholars divide actions into two separate classes, one for conflictual actions and another for cooperational. Next, they calculate the mean or sum score for conflictual and coopera-

[10] Because the IPI scale only contains positive scores, the following discussion of sum and mean scales is not applicable to the IPI data.

tional events separately.[11] Although this is more logical than taking the sum or mean of all events together, it is still vulnerable to the sum problem and the mean problem defined above.

2.2 Counts

In an attempt to best overcome the three problems above, a number of studies utilize count data. "net cooperation" is the most straightforward count, which is simply the total number of cooperational actions - number of conflictual actions according to the Goldstein scale. in order to attempt for most variation in the types of events that are occurring, a number of other scholars utilize Duvall and Thompson counts Duval and Thompson (1980), which place all events on the WEIS scale into the four conceptually unique categories below.[12]

- *Verbal Cooperation*: The occurrence of dialogue-based meetings (e.g. negotiations, peace talks), statements that express a desire to cooperate or appeal for assistance (other than material aid) from other actors.
- *Material Cooperation*: Physical acts of collaboration or assistance, including receiving or sending aid, reducing bans and sentencing, etc.
- *Verbal Conflict*: A spoken criticism, threat, or accusation, often related to past or future potential acts of material conflict.
- *Material Conflict*: Physical acts of a conflictual nature, including armed attacks, destruction of property, assassination, etc.

Below, I provide a list that illustrates the type of action aggregation strategy utilized in a selection of prominent event data studies.

- **Scale**
  - Goldstein mean
    - Sky (2000); Fordham (2005); Brandt and Freeman (2004); Shellman (2004*b*); Shellman and Stewart (2007); Shellman, Hatfield and Mills (2010); Shellman (2007); Brandt, Colaresi and Freeman (2008)
  - Goldstein sum
    - Schrodt and Gerner (1997, 2000); Goldstein et al. (2001); Stoll and Subramanian (2006); Schrodt and Gerner (2001); Gleditsch and Beardsley (2004); Schneider and Troeger (2006); Shellman, Hatfield and Mills (2010)
  - IPI scale
    - Shellman (2004*a*,*b*)
  - IDEA scale
    - Hämmerli, Gattiker and Weyermann (2006), Bond et al. (2003)
  - VICDP scale
    - Moore (1998, 2000)

[11] Unlike Goldstein and IDEA scale, the IPI action typology divides conflictual and cooperational events. Consequentially, scholars choosing to scale or sum use IPI-coded event data must do so for conflictual and cooperational events separately.

[12] This approach is easily extendible to eh CAMEO coding scheme since it is built on WEIS' general framework.

- **Counts**
  - Goldstein counts (positive and negative)
    - Pevehouse (2004); Shellman, Hatfield and Mills (2010)
  - Net cooperation
    - Goldstein (1991, 1997)
  - Duvall and Thompson counts
    - Brandt, Freeman and Schrodt (2011) (only material conflict events), Shearer (2006), Schrodt, and Yilmaz (2003) (adds a fifth category called "Mediation and negotiation"), Schrodt (2006); D'Orazio, Yonamine and Schrodt (2011)
- **Action codes**
  - Kovar et al. (2000)
  - Schrodt (2006)
- **Other**
  - Bond et al. (1997) – Conflict carrying capacity

3 Temporal Aggregation

After completing the first two steps of the aggregation process, researchers must determine how to account for time. The vast majority of studies temporally aggregate event data into a man-made length of time, commonly by day, week, month, quarter, or year. The choice of temporal unit determines that length of time across which the action aggregation is performed. For example, consider a scholar who is interested in measuring the Goldstein sum of events occurring between the Indian military and Indian rebels. The actor and action aggregation steps (i.e. Section 1 and Section 2) would result in a dataset containing only events that occurred between the actors of interest with an extra column indicating the Goldstein score for each action. To complete the aggregation process, the researcher must determine a temporal unit of aggregation. If the researcher chooses the weekly level, he/she would simply calculated the sum of the Goldstein scores for events occurring in the same week.

As mentioned in the introduction to this study, a number of articles demonstrate that different temporal aggregations can affect the empirical results (see Freeman (1989); Alt, King and Signorino (2001); Dale (2002); Shellman (2004*a*)). These analyses suggest that it is important to both theoretically justify the level of aggregation and, if possible, employ multiple levels as robustness checks. Below, I list the temporal aggregations employed in a number of prominent studies.[13]

Although most scholars aggregate temporally, a small body of studies focus on the sequential order of events, irrespective of traditional units of time. Drawing on Marlin-Bennett, Rosenblatt and Wang (1991), Moore (1998)

[13]Note that certain studies, like Shellman (2004*a*) and D'Orazio, Yonamine and Schrodt (2011) follow best practice and perform analyses across different levels of temporal aggregation to serve as robustness.

builds event sequence based on "moves" and "turns" as opposed to standard temporal units.[14] Subsequence work, including Moore (2000); Shellman (2004*a*, 2007) adopt a similar "move" and "turn" approach to model event data. Likewise, some studies using HMMs do not confine event patterns to traditional temporal units.[15] For example, Schrodt (2000) focuses on identifying and predicting transitions between distinct phases of conflict irrespective of the amount of time that a state exists in a given phase.

Below, I illustrate the temporal aggregation choices employed within the event data literature.

- Daily
  - Goldstein (1991); Pevehouse and Goldstein (1999); Goldstein et al. (2001); Shellman (2004*a*,*b*); Schrodt (2006); Schneider and Troeger (2006); Shearer (2006)
- Weekly
  - Goldstein (1991); Goldstein et al. (2001); Brandt and Freeman (2005); Shellman and Stewart (2007); Brandt and Freeman (2004); Shellman, Hatfield and Mills (2010); Goldstein (1997); D'Orazio, Yonamine and Schrodt (2011)
- Bi-weekly
  - Stoll and Subramanian (2006)
- Monthly
  - Goldstein (1991); Schrodt (1997); Sky (2000); Schrodt, and Yilmaz (2003); Schrodt (2007); Schrodt and Gerner (1997, 2000, 2001); Shellman (2004*a*,*b*); Ward, Greenhill and Bakke (2010); Gleditsch and Beardsley (2004); Shellman, Hatfield and Mills (2010); Brandt, Freeman and Schrodt (2011); D'Orazio, Yonamine and Schrodt (2011); Brandt, Colaresi and Freeman (2008)
- Quarterly
  - Jenkins and Bond (2001); Fordham (2005); Shellman (2004*a*)
- Yearly
  - Pevehouse (2004); Bond et al. (1997)
- Non-temporally defined sequences
  - "Move" and "Turn" sequences
    - Moore (1998, 2000); Shellman (2007, 2004*b*)
  - Markov transition sequences
    - Schrodt (2000)
- Unspecified
  - Hämmerli, Gattiker and Weyermann (2006)

[14] For more information on the rationale behind analyzing events in terms of sequences, see Schrodt (2000).

[15] This is not true of all event data studies using HMMs. For example, Shearer (2006) applies HMMs to temporally aggregated data to facilitate the substantive interpretation of his findings.

4 Conclusion

Currently, we are experiencing a number of trends that are altering the nature of data in political science, two of which are of particular relevance to this study. First, the amount of open source electronic text – from various outlets ranging from traditional news media to social networking sites – is growing at an increasing rate. Second, advancements in machine-coding mean that we are able to extract more detailed and accurate information form text faster than ever. These two developments, as well as an increasing focus in the conflict literature at geo-spatially and temporally nuanced analyses, suggests that event data will become increasingly prominent moving forward. Given this, it is crucial that scholars take their actor, action, and temporal aggregation decisions seriously. Perhaps even more important, scholars should carefully detail and justify these decisions – either theoretically, empirically, or with a combination of the two – in their research designs.

Many other areas of political science have established clearly defined coding rules. For example, the Correlates of War provide clear a clear definition how many specific actions (battle fatalities) must occur between specific actors (government forces and a domestic rebel group) within a set period of time (one calendar year) to qualify as an "intra-state war". Despite the massive number of aggregation options and proven effect that even minor changes can have on empirical results, it is surprising that studies utilizing event data have yet to adopt a formal framework for selecting and documenting aggregation choices.

In the In this paper, I provide a general overview of the aggregation techniques present in the extant literature, highlighting relevant strengths and weaknesses of the various approaches and "best practices" when they exist. I hope that this facilitates scholars' future decision making process regarding aggregation choices in terms of actors, actions, and temporal units and encourages thorough documentation.

References

Alt, James, Gary King and Curt S. Signorino. 2001. "Aggregation among Binary, Count, and Duration Models: Estimating the Same Quantities from Different Levels of Data." *Political Analysis* 9:21–44.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor and Kurt Schock. 1997. "Mapping Mass Political Conflict and Civil Society: Issues and Prospects for the Automated Development of Event Data." *Journal of Conflict Resolution* 41(4):553–579.

Bond, Doug, Joe Bond, Churl Oh, J. Craig Jenkins and Charles L. Taylor. 2003. "Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development." *Journal of Peace Research* 40(6):733–745.

Bond, Joe, Vladimir Petroff, Sean O'Brien and Doug Bond. 2004. "Forecasting Turmoil in Indonesia: An Application of Hidden Markov Models." Presented at the International Studies Association Meetings, Montreal.

Brandt, Patrick T. and John R. Freeman. 2004. "Testing Democratic Peace Theory: A New Approach with Application to the Israeli-Palestinian Conflict." Presented at the International Studies Association Meetings, Montreal.

Brandt, Patrick T. and John R. Freeman. 2005. "Advances in Baysian time Series Modeling and the Study of Politics: Theory testing, Forecasting, and Policy Analysis." *Political Analysis* 14:1–36.

Brandt, Patrick T., John R. Freeman and Philip Schrodt. 2011. "Real Time, Time Series Forecasting of Inter- and intra-State Political Conflict." *Conflict Management and Peace Science* 28(1):41–64.

Brandt, Patrick T., Michael P. Colaresi and John R. Freeman. 2008. "The Dynamics of Reciprocity, Accountability and Credibility." *Journal of Conflict Resolution* 52(3):343–374.

Dale, Thomas G. 2002. "Event Data Analysis and Threats from Temporal Aggregation." Presented at the Florida Political Science Association Meeting, Sarasota.

D'Orazio, Vito, James E. Yonamine and Philip A. Schrodt. 2011. "Predicting Intra-state Conflict Onset: An Event Data Approach Using Euclidean and Levenshtein Distance Measures." Presented at the annual Midwest Political Science Association meeting, Chicago.

Duval, Robert D. and William R. Thompson. 1980. "Reconsidering the Aggregate Relationship between Size, Economic Development, and Some Types of Foreign Policy Behavior." *American Journal of Political Science* 24(3):511–525.

Fordham, Benjamin O. 2005. "Strategic Conflict Avoidance and the Diversionary Use of Force." *The Journal of Politics* 67(1):pp. 132–153.

Freeman, John R. 1989. "Systematic Sampling, Temporal Aggregation, and the Study of Political Relationships." *Political Analysis* 1:61–98.

Gleditsch, Kristian Skrede and Kyle Beardsley. 2004. "Noisy Neighbors: Third-Party Actors in Central American Conflicts." *Journal of Conflict Resolution* 48(3):91–119.

Goldstein, Joshua S. 1991. "Reciprocity in Superpower Relations: An Empirical Analysis." *Journal of Conflict Resolution* 36:369–385.

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

Goldstein, Joshua S. 1997. "Reciprocity in Superpower Relations: An Empirical Analysis." *International Studies Quarterly* 35(2):512–529.

Goldstein, Joshua S., Jon C. Pevehouse, Deborah J. Gerner and Shibley Telhami. 2001. "Reciprocity, Triangularity, and Cooperation in the Middle East, 1979-1997." *Journal of Conflict Resolution* 45(5):594–620.

Hämmerli, August, Regula Gattiker and Reto Weyermann. 2006. "Conflict and Cooperation in an Actor's Network of Chechnya based on Event Data." *Journal of Conflict Resolution* 50(159):159–175.

Jenkins, Craig J. and Doug Bond. 2001. "Conflict Carrying Capacity, Political Crisis, and Reconstruction." *Journal of Conflict Resolution* 45(1):3–31.

Kovar, Klaus, Johannes Petrak Fürkranz, Johann Petrak, Bernhard Pfahringer, Robert Trappl and Gerhard Widmer. 2000. "Searching for Patterns in Political event Sequences: Experiments with the KEDS Database." *Cybernetics and Systems* 31(6).

Marlin-Bennett, Reneé, Alan Rosenblatt and Jianxin Wang. 1991. "The Visible Hand: The United States, Japan, and the Management of Trade Disputes." *International Interactions* 17:191–213.

Moore, Will H. 1998. "Repression and Dissent: Substitution, Context, and Timing." *American Journal of Political Science* 42(3):851–873.

Moore, Will H. 2000. "The Repression of Dissent: A Substitution model of government coercion." *Journal of Conflict Resolution* 44:107–127.

Moore, Will H. and Ronny Lindström. 1996. "The Violent Intranational Conflict Data Project (VICDP) Codebook." University of California Riverside, typescript.

Pevehouse, Jon C. 2004. "Interdependence Theory and the Measurement of International Conflict." *Journal of Politics* 66(1):247–266.

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

Schneider, Gerald and Vera E. Troeger. 2006. "War and the World Economy: Stock Market Reactions to International Conflict." *Journal of Conflict Resolution* 50(5):623–645.

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

Schrodt, Philip A. 2000. Pattern Recognition of International Crises using Hidden Markov Models. In *Political Complexity: Nonlinear Models of Politics*, ed. Diana Richards. Ann Arbor: University of Michigan Press pp. 296–328.

Schrodt, Philip A. 2006. Forecasting Conflict in the Balkans using Hidden Markov Models. In *Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention*, ed. Robert Trappl. Dordrecht, Netherlands: Kluwer Academic Publishers pp. 161–184.

Schrodt, Philip A., and Deborah J. Yilmaz, Omur Gerner. 2003. "Evaluating "Ripeness" and "Hurting Stalemate" in mediated International Conflicts: An Event Data Study of the Middle East, Balkns, and West Africa." Presented at the Annual Meeting of the International Studies Association, Portland, Oregon.

Schrodt, Philip A. and Deborah J. Gerner. 1997. "Empirical Indicators of Crisis Phase in the Middle East, 1979-1995." *Journal of Conflict Resolution* 25(4):803–817.

Schrodt, Philip A. and Deborah J. Gerner. 2000. "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94(4):803–817.

Schrodt, Philip A. and Deborah J. Gerner. 2001. "Analyzing the Dynamics of International Mediation Processes in the Middle East and the Former Yugoslavia." Presented at the annual meeting of the International Studies Association, Chicago.

Schrodt, Phillip A. 2007. "Inductive Event Data Scaling using Item Response Theory." Presented at the Summer Meeting of the Society of Political Methodology. Available at http://eventdata.psu.edu.

Shearer, Robert. 2006. "Forecasting Israeli-Palestinian Conflict with Hidden Markov Models." Available at http://eventdata.psu.edu/papers.dir/Shearer.IP.pdf.

Shellman, Stephen. 2004*a*. "Time Series Intervals and Statistical Inference: The Effects of Temporal Aggregation on Event Data Analysis." *Political Analysis* 12(1):97–104.

Shellman, Stephen. 2007. "Process Matters: Conflict and Cooperation in Sequential Government-Dissident Interactions." *Security Studies* 15(4):563–599.

Shellman, Stephen and Brandon Stewart. 2007. "Political Persecution or Economic Deprivation? A time-Series Analysis of Haitian Exodus, 1990-2004." *Conflict Management and Peace Science* 24:121–137.

Shellman, Stephen, Clare Hatfield and Maggie Mills. 2010. "Dissagregating Actors in Intrastate Conflict." *Journal of Peace Research* 47(1).

Shellman, Stephen M. 2004*b*. "Measuring the Intensity of International Political Interactions Event Data: Two Interval-Like Scales." *International Interactions* 30(2):109–141.

Sky, David J. 2000. "Routine, Response, or Rational Expectations?: An Analysis of Bilateral Relations in Intranational and International Contexts." Presented at the Annual Meeting of the International Studies Association, Los Angeles, CA.

Stoll, Richard J. and Devika Subramanian. 2006. "Hubs, Authorities, and Networks: Predicting Conflict using Events Data." Presented at the Annual Meeting of the International Studies Association, San Diego, CA.

Thomas, Dale G. 2002. "Event Data Analysis and Threats from Temporal Aggregation." Presented at the Florida Political Science Association Meeting, Sarasota, Fl.

Vincent, Jack E. 1983. "WEIS vs. COPDAB: Correspondence Problems." *International Studies Quarterly* 27:161–168.

Ward, Michael D., Brian D. Greenhill and Kristin M. Bakke. 2010. "The Perils of Policy by P-Value: Predicting Civil Conflicts." *Journal of Peace Research* 47(5).