ANALYSIS OF EVENTS IN THE STUDY OF COLLECTIVE ACTION

Susan Olzak

Department of Sociology, Uris Hall, Cornell University, Ithaca, New York 14853

Abstract
Recent research on collective action has focused on the occurrence, timing, and sequencing of such events as regime changes, riots, revolutions, protests, and the founding of social movement organizations. Event analysis allows information on the duration, number of participants, presence of violence, or outcome of some particular type of collective action to be compared across social systems or across time periods. This review considers issues of definition, measurement, and methods of estimation in event analysis. It also compares two general varieties of event analysis: approaches that model the dynamics of collective action as a process, and those that do not. A process-oriented approach evaluates how time and covariates (including past events) affect the timing and sequence of repeatable events, and it attempts to explain how events unfold over time. The nonprocess approaches summarize static relationships between levels or characteristics of units and some type of event count.

INTRODUCTION
Sociologists interested in the causes of social movements and their success or failure have long considered natural histories of movements and revolutions, profiles of activists, and catalogues of grievances. Recently, spurred mainly by Charles Tilly’s pathbreaking research, they have begun to focus on the rate of occurrence of collective events. This chapter reviews and evaluates methods for using information on events—especially their timing and sequences—to analyze social movements and collective action.1

1See Jenkins (1983), Marwell & Oliver (1984), Zald & McCarthy (1979), and Tarrow (1988a) for recent substantive reviews of theory and research on collective action.
Event analysis has an extremely broad scope. Studies of conflicts have included race riots (Spilerman 1970), strikes and industrial protests (Aminzade 1984), political violence (Tilly, Tilly & Tilly 1975), peasant rebellions (Paige 1975), revolutionary activity (Markoff 1986), lynchings (Inverarity 1976), and coups d’etat (Hannan & Carroll 1981). Studies of protest have considered black civil rights’ protests (McAdam 1982), marches against nuclear power plants (Walsh & Warland 1983), Québécois separatist demonstrations (Olzak 1982), union organizing of farm workers (Jenkins & Perrow 1977), protests by American Indians (Nagel 1988), and social protests in Italy (Tarrow 1988b). Studies of foundings and disbandings of social movement organizations have included such varied assemblies as labor unions (Hannan & Freeman 1987), anti-drunk driving groups (McCarthy et al 1988), and civil rights’ organizations (Jenkins & Eckert 1986).

Event analysis allows diverse forms of collective action to be measured and compared because observations are collected in commensurate dimensions. Thus, information about revolutions consisting of thousands of acts of violence and about one-time confrontations is enumerated and described in terms of number of participants, duration of unrest, magnitude of violence, and other characteristics. This feature of the strategy is important because it permits cases of unsuccessful and/or short-lived movements to be compared with instances of successful and long-lived ones. It also allows stages that emerge during collective actions to be analyzed as a sequence of different kinds of events rather than as one continuous action. In short, the strategy of studying events uses more information about the dynamics of change in social movements than do strategies that treat movements as unitary phenomena.

THE DEVELOPMENT OF EVENT ANALYSIS

Event analysis took shape during the 1960s in research on variation among countries in collective violence and on causes of race riots. Subsequent advances in methods for analyzing historical and time-ordered data on events have increased the power of event analysis and encouraged research on classic questions about the influence of timing and organizational dynamics on rates of collective action and on the success of social movements.

Brief History of Event Analysis

Cross-national research on collective violence in the 1960s related annual counts of events to economic, political, and other structural variables to learn whether peak periods of economic hardship were peak periods of unrest. These studies developed formalized rules for coding information on collective events using records from archives, newspapers, historical documents, and police or magistrate records (for a recent example, see Taylor & Jodice 1983).
They also produced data banks that allowed researchers to compare effects of structural and political variables on classes of events among nations and over historical periods (Rule & Tilly 1965, Feierabend & Feierabend 1966, Feierabend et al 1972). During this period, research moved beyond describing and categorizing stages in conflicts to analyze causes of unrest. This shift had two important consequences. First, use of event counts opened up the possibility that diverse social science methods could be used to analyze questions about social movements. Second, the substantive findings informed theoretical debates about social upheavals. In particular, theories of social protest that highlighted mechanisms of relative deprivation and social disorganization unraveled once they were subjected to empirical scrutiny (Snyder & Tilly 1972, Rule & Tilly 1972, Tilly, Tilly & Tilly 1975, Tilly 1981, Weede 1987).

A concurrent line of research related demographic and socioeconomic characteristics of communities to counts of race riots. This research also informed theoretical debates about civil unrest. The findings challenged the assumption that communities experiencing greatest economic hardship, poverty, and relative and absolute deprivation were most likely to experience race riots (Lieberson & Silverman 1965, Wanderer 1969, Spilerman 1976, Eisinger 1973, McPhail & Wohlstein 1983).

Motivations for Event Analysis

Choice of a method for analyzing data on collective action depends upon a researcher’s purposes. If the goals involve testing generalizable theories with replicable evidence, then event analysis has several potential advantages. First, as noted above, event analysis can use statistical techniques and models that have been useful in other subfields of sociology and other sciences. Hypotheses can be tested and models can be compared on their relative merits (Spilerman 1971, Paige 1975). Although case studies provide rich detail and historical background, they rarely provide systematic evidence appropriate for testing hypotheses. For example, social scientists commonly study only the places that have collective events in studies that try to learn the causes of events. For instance, Wanderer’s (1969) analysis of riot severity used data only for cities that had race riots, thereby ignoring characteristics of cities that did not have a riot and sidestepping the question of why riots occurred. Sampling on the dependent variable, as in these examples, confounds causal interpretation because (by definition) some observations have been excluded from analysis based upon their levels on the dependent variable (success or not, participation or not). Any attempt to infer causal relationships will be crippled by sample selection bias (Heckman 1979, Berk 1983). Sample selection bias has often been confused with the issue of representativeness of samples. They are not the same. For example, data on events occurring in
Utah may be questioned on the grounds that the sample is not a random one and because Utah is not representative of other geographical regions. If the goal is to estimate some mean number of events in the country as a whole, then using this sample will cause problems. If the goal is to estimate the effects of some covariates on the rate of activity, the estimates from such a sample may be unbiased. This is because nonrepresentativeness hampers one’s ability to estimate population means and variances but does not necessarily affect estimates of structural parameters, e.g. effects of covariates on the rates. Sample selection bias on the other hand means that a sample has been chosen based upon some value of the dependent variable, and effects of covariates on the rates will be biased unless appropriate corrections are applied (Heckman 1979).

An important motivation for event analysis is thus to avoid sample selection bias. Research designs in event analysis avoid problems of sample selection bias by including both successful and unsuccessful movements (Gamson 1975), cities with and without racial conflicts (Spilerman 1970), and samples drawn from populations of participants and nonparticipants (McAdam 1986).

Furthermore, case studies of social movements normally do not exploit information on timing and sequences of events. One important form of event analysis, event-history analysis, does analyze such information. This is particularly important in studying processes of temporal contagion in the spread of riots, rebellions, and other disturbances.

Some have suggested that the spread of unrest by geographical contagion may also be important. In one case, consideration of spatial contagion altered substantive conclusions: Doreian (1981) found that spatial autocorrelation affected estimates of factors causing local variations in participation in the Huk rebellion in the Philippines. Investigating the intersection of social geography and diffusion holds promise for examining how proximity affects contagion (Markoff 1985).

Whether the occurrence of collective actions means that a social movement exists is treated as problematic and open to empirical examination in event analysis. This stance anticipates the possibility that events might occur after a revolution has ended or before a social movement supposedly begins. This strategy stands in contrast to traditional ones that analyze social movements only within conventional temporal boundaries defined by the history of some movement or revolution.

Restricting research to a priori temporal units of revolution, civil war, or social movements may miss much variation in activity over time and over regions. For instance, Tilly, Tilly, & Tilly (1975: Figure 6) show that the number of collective disturbances in France has four peaks during 1830–1930: disruptions of the revolution of 1830, the unrest surrounding the revolutionary activity in 1848, strikes around 1906, and social unrest during the mid-1930s.
What surprised even the authors was that the four peaks contained comparable numbers of violent events. Apparently other “revolutions” of equal significance occurred in addition to that in 1830.

Event analysis also allows empirical investigation of changes in forms of collective action. For example, Tilly’s (1986) analysis of French contentious gatherings reveals that collective actions were transformed over a hundred years from local to national scope. This approach also allows forms of conflict to be analyzed separately by target and degree of violence, facilitating examination of similarities and differences in effects of covariates on types of events. In other words, claims that causes of various types of collective action differ can be examined empirically, rather than assumed. Causal links between forms of collective action can also be examined. For instance, there has been much speculation but little evidence about the link between strikes and racial violence; some suggest that the occurrence of strikes (and the use of black strikebreakers) led to racial violence. Olzak (1989) found that growth of labor unions—but not strike frequency—raised rates of violence against blacks but decreased rates of violence against white immigrants in late nineteenth and early twentieth century America.

Another source of the increasing popularity of event analysis is its flexibility in using different levels of spatial aggregation. Events have been studied as products of groups, neighborhoods, electoral districts, export sectors, and nation-states. Consider the case of American race riots during the 1960s. Stark et al (1974) focused on geographical spread during one year of race rioting over census tracts of one city, Los Angeles. Spilerman (1970) analyzed data on the annual occurrence of riots in numerous cities in order to learn whether characteristics of cities affect the probability of a riot. Still others analyzed yearly counts of riots for the United States as a whole in order to learn whether they fluctuated with variation in the national political and economic environment (Kelly & Isaac 1984).

**Promising Developments in Event Analysis**

Researchers have built on the studies of the 1960s and have expanded the repertoire of methods for collecting and analyzing events. Two areas deserve special note. The first concerns life histories of social movement organizations (SMOs). Gamson (1975) and McCarthy & Zald (1977) proposed a research agenda that emphasized the importance of gathering information on such organizations. But this agenda was not quickly implemented. Until recently only Steedly & Foley (1979) and Goldstone (1980) analyzed SMOs directly (and both used Gamson’s data). This lack of attention to organizations is curious given the centrality of organizational structure in resource mobilization theories of collective action. However, the situation seems to be changing. Hannan & Freeman (1987, 1989), McCarthy et al (1988), and
Carroll & Huo (1988) have studied the founding and disbanding rates of SMOs, and Rosenthal et al (1985) studied network relationships among organizations in the women's movement.

A second promising line of research investigates "cycles" or waves of protest empirically. Beginning with Durkheim's analyses of social disorganization, the idea that social movements come in recurrent patterns has intrigued sociologists. Although waves, cycles, and episodes serve as core concepts in theories of collective behavior (Turner & Killian 1957), few researchers have studied such patterns. Hamblin et al (1973) suggested that mathematical models of contagion could be applied to analysis of diffusion of collective events. Only Spilerman (1971) and Olzak (1987a) have modeled processes of contagion directly.

Both developments—analysis of organizational events and analysis of trajectories of diffusion—highlight the importance of gathering temporal data. Theory and research now attend to the effects of the histories of social movements (and their internal organizational developments) and of sequences of protests and conflicts on the likelihood of future events.

DEFINITIONS IN EVENT ANALYSIS

Conceptual Definitions

This research tradition commonly defines events as nonroutine, collective, and public acts that involve claims on behalf of a larger collective (Paige 1975, Tilly 1978). Thus a minimal definition of collective action is that it (a) involves more than one person, and (b) makes claims of agency (or corporate) status. Tilly's (1978) list of claims includes petitioning, memorializing, and opposition to or support of an enemy of the government. Tilly (1978: 275) defines agency status as involving acts that "make a visible claim which, if realized, would affect the interests of some specific person(s) or group(s) outside their own number."

Should social movements and events be defined independently? If so, what should be the distinguishing criteria? According to the minimal definition, social movements and collective actions are separable. Social movements are "a group of people identified by their attachment to some particular set of beliefs" (Tilly 1978: 9). Conceptual definitions of collective actions normally do not require the presence of shared beliefs among actors. However, some researchers stipulate that collective actions involve groups with preexisting solidarity (e.g. Paige 1975: 90). Others implicitly assume that solidarity is

---

2In the mid-1970s there was probably less consensus about what distinguished political collective actions from panics, fads, etc. To test theories about political collective actions, some distinctions among types of collective behavior were necessary. Thus Paige's motivation for including solidarity was therefore to exclude collective behaviors such as "panics, mass movements of refugees, crazes . . ." (1975: 87, 390).
necessary for the types of collective actions they have analyzed, but leave open the issue of whether solidarity is necessary in other instances.

Some researchers also stipulate that collective action be noninstitutional (Paige 1975), while others take this position implicitly by limiting analysis to actions instigated by challenging groups that are "countercultural" or that have little formal power (Kriesi 1988). Perhaps because it implies a process, institutionalization has proved to be difficult to operationalize. Groups that have already gained acceptance in the political arena are thus often excluded from analysis of social movements and collective action. Lowi (1971) provided a rationale for this practice, claiming that SMOs that succeed in establishing routinized and stable ties with a government agency have been transformed into interest groups. In this view, strikes by most labor unions are now actions of interest groups, not social movements (Snyder 1975). Using Lowi's criteria, actions of some "countermovements" could be excluded, despite the fact that research on them has illuminated ways that members of groups with access to power sometimes form significant social movements. Examples include studies of the pronuclear power movement (Useem & Zald 1982), the antiabortion movement (McCarthy 1987), and coups d'état by dissatisfied army units (e.g. Feierabend & Feierabend 1966).

Tilly takes a different tack. He defines collective action to potentially include actions of all contenders, including both "members," who are those who have "routine, low-cost access to resources controlled by the government," and "challengers," who are other contenders (Tilly 1978: 52). If, as Tilly suggests, both members of the polity and outsiders can conduct social movements, then we should analyze transformations of outsiders into members of the polity as events in the life histories of social movements. Such analysis requires attention to the boundaries that separate insiders from outsiders.

But there is disagreement about whether institutionalization of social movements belongs in the study of collective action. Some researchers argue for excluding institutional behavior because it concerns elite activity rather than spontaneous grass-roots mobilization (Jenkins & Perrow 1977, McAdam 1982). Others think that it is unreasonable to truncate observations on SMOs at the point when they gain access to the political arena. Such exclusion assumes that institutionalization is a one-way process and ignores the fact that some established groups become delegitimated and excluded from the polity. Excluding institutional action from studies of collective action by definition makes it impossible to address a fundamental sociological process: how charismatic movements become routinized and institutionalized over time.

If we add institutional activities to the list of potential collective actions, we face the prospect that the concept encompasses virtually all social behavior including everyday actions such as business conferences, professional meetings, and church assemblies. Tilly's (1978) emphasis on excluding routine
activity has been a key solution to reduce the scope. There are often good reasons to exclude anniversary marches or annual celebrations since their timing is predetermined and not spontaneous. Researchers also usually exclude routine congressional activities, court sessions, political party conventions, and regular government activity from the concept of collective action (Burstein 1985). According to Tilly's guideline, routine church activities would not be coded as events, but nonroutine actions of church members participating in civil rights' marches under the leadership of a minister might be.

The minimalist definition of collective events has advantages over definitions that include dimensions such as size, noninstitutional status, violence, and political orientation as defining features of collective action. Relaxing the requirement that collective action involves self-conscious identity, solidarity, or shared sentiment allows researchers to examine whether solidarity affects rates of collective action. In contrast, enumerating only events that reflect solidarity vitiates the possibility of studying the effects of variations in solidarity. It renders tautological the key claim of resource-mobilization theory that solidarity affects collective action. So it appears that a sensible strategy in event analysis is to begin with minimal definitions of events so that the effects of solidarity, institutionalization, and other dimensions can be examined empirically.

Use of the minimal definition has a disadvantage. Coding costs increase with the scope of relevant actions. In addition, effort is needed to narrow the focus of the research later rather than earlier in the research process. But in many cases, the capacity to narrow the scope, or to compare different forms of events, after the data are collected may outweigh these costs.

**Operational Definitions**

At what size does an event becomes collective? The answer varies widely among studies and types of event. Most researchers choose a minimum in the range of 10 and 20; Tilly (1978) proposed a minimum of 50, and Olzak and West (1987) use a minimum of 2. But sources of data often lack information on exact numbers. Olzak (1987) found that daily accounts of ethnic protests in the New York Times provide reliable counts of participants for only about half of the events. Fortunately, newspaper reports nearly always contain key phrases such as "small group of hooligans" and "large crowd" that indicate approximate size categories.

Most researchers hold that events occur in one local setting, although some Civil Rights campaigns and intercity marches are exceptions to this rule (Morris 1984). But geographical units change over time, and information is sometimes found in incompatible units. Markoff & Shapiro (1973) discuss solutions for problems that occur when geographical units change over
COLLECTIVE EVENT ANALYSIS

time because boundaries of census, political, electoral, and other kinds of districts change.

Decisions about temporal boundaries of events are especially complicated. When does one event stop and a new one begin when events are closely spaced in time? Paige (1975), Tilly (1978), and Olzak (1987) judge a new event to have begun if activity resumes after a pause of at least 24 hours in a locale. Otherwise, activity of more or less continuous actions by the same group is considered as one event.

Marwell & Oliver (1984) argue that "collective campaigns" are more useful units than are events. Certainly the concept of a campaign can be useful, as in the case of the Civil Rights' Movement in which marches carried across state boundaries. However, the concept of campaign also faces serious problems. When does an event become a campaign? This would be a problem for an abortive campaign or an isolated event that might have begun a campaign but failed. And, when does a campaign end? Just as smaller geographical units are more easily aggregated to larger entities than vice-versa (Markoff & Shapiro 1973), a strategy of analysis (rather than measurement) that specifies how to link events that make up a campaign may be more useful than one that begins with observations on campaigns. Moreover, collapsing brief episodes of many events into one campaign makes the occurrence of many small events indistinguishable from one uninterrupted protest. Such a coding strategy makes it difficult to discover cycles of protests.

It is often important to measure variation in intensity or magnitude of events. Some researchers combine information on numbers of participants and duration of events in order to measure intensity. One common approach weights events by their duration (in days) multiplied by the average number of participants per day, thereby measuring events in so-called person-days (Rule & Tilly 1965, Hibbs 1976, Olzak 1982, Rasler 1986).

Much research on collective action focuses on collective violence. Exactly what constitutes a violent event? And whose violence provides the benchmark—protesters, police, or outsiders, or all three? Again operationalizations vary widely. A violent event has been defined as "an instance of mutual and collective coercion within an autonomous political system which seizes or physically damages persons or object? (Tilly 1978: 248). Others treat events as violent if weapons or attacks are used merely to threaten property or persons (Olzak 1989). Political scientists have measured political violence in terms of the number of deaths from political violence (Muller 1985, Weede 1987). Categories used in coding violence and size are sometimes interdependent. In some coding schemes, occurrence of violence affects whether or not an event is taken regardless of size. For example, all violent events are included in some studies of contentious gatherings and protests regardless of size (Tilly 1978, Tarrow 1983).
DATA AND MEASUREMENT

Those who analyze quantitative data on events sometimes use data and methods without questioning their drawbacks and limitations (Franzosi 1987). This section reviews some disadvantages, problems, and constraints encountered in event analysis and some proposed solutions.

Sources and Methods of Data Collection

Collection of information on date of occurrence, duration, participants, targets, and violence relies increasingly on:

a. official archival records, such as police records and municipal records of arrests and/or deaths from civil unrest (Tilly, Tilly, & Tilly 1975);

b. annual newspaper indexes (Spillman 1970, McAdam 1982, Burstein 1985, Jenkins & Eckert 1986);

c. daily newspaper accounts (Paige 1975);

d. published listings of events originally compiled from various newspaper sources, e.g. the widely used data on strikes in the United States (Griffin 1939, US Bureau of the Census 1975), lynchings (NAACP 1919, US Bureau of the Census 1975), and political turmoil (Taylor & Jodice 1983); and

e. organizational histories in newsletters or other secondary historical accounts (Griffin 1939, Hannan & Freeman 1987, 1989, McCarthy et al 1988).

Much research reviewed here has come to rely primarily on newspaper reports in part because such sources can be followed forward in time and because at least some researchers have found that newspapers provide the most complete account of events for the widest sample of geographical or temporal units (Tilly, Tilly, & Tilly 1975: 16). Research on the United States has treated the New York Times (NYT) as a source for the entire country and much of the debate has focused on this source.

Reliance on newspapers as a primary source of data has its critics. The most common criticism leveled at this approach is the suspicion that the data contain some systematic biases (Johnson et al 1971, Chermesh 1982, Franzosi 1987). For example, Danzger (1975) finds that the location of UPI and AP wire offices affected coverage of riots in the NYT. Kielbowicz & Scherer (1986) claim that more violent, larger, or longer events are more likely to be covered in newspapers and that coverage affects subsequent events. Unfortunately there is rarely a way to evaluate these claims directly since few alternative sources contain as much information. However, Snyder & Kelly’s (1977) analysis comparing coverage in local and national newspapers shed
light on these issues. They compared accounts of riots in daily NYT accounts with those from local newspapers in 43 American cities from 1965 to 1969. They found that the location of offices of major wire services was not related to the likelihood that an event in a local newspaper would be reported in the NYT but that conflict intensity, measured in terms of size, violence, and duration of an event did. So Snyder & Kelly (1977) recommended that some threshold of the number of participants in events should be used, or that events of similar intensity should be analyzed separately when using data from the NYT. For problems related to geographical bias that arise when, for instance, only the NYT is used, Franzosi (1987) suggests that some alternative source—preferably a local newspaper—be used as a supplement. However, the NYT remains the most complete source of data on events, according to analyses that compared coverage in the NYT with the Los Angeles Times and other west coast newspapers, for data on collective actions that occurred only on the west coast (Jenkins & Perrow 1977).

**Instruments for Collecting Data on Events**

Charles Tilly and his associates remain the leaders in providing methods and techniques for collecting information on events of political and contentious actions (Rule & Tilly 1965, Tilly et al 1975, Tilly 1978, Tilly 1986, Horn & Tilly 1986). Others have adapted his approach for other kinds of events. Especially detailed coding appendices and technical reports include those by Gamson (1975), Tilly (1978), McAdam (1982), Paige (1975), Tarrow (1981), Burstein (1985), Horn & Tilly (1986), and Olzak & West (1987). Compendia of events in various countries include those by Botz (1976), Kriesi (1981), Sugimoto (1981), and Horn & Tilly (1986).

Although studies differ in defining event-forms, duration, violence, size, and outcomes, two main differences in strategy deserve note. First, Horn & Tilly (1986) and Franzosi (in press) developed computer programs that analyze the content of newspaper data on events. These programs allow researchers to apply methods of content analysis to original newspaper accounts of strikes, revolutions, and racial strife using consistent rules. There are trade-offs involved in using open-ended programs, however. While they allow analysis of actual word counts, nuances, and perhaps also biases of particular reporters or specific newspapers (Franzosi 1987), such programs require a large investment in coding time and computer space. Moreover, using content analysis means that a second stage of distilling word counts and evaluation is necessary before beginning event analysis. Using shorter and simpler "coding sheets" that record data in a fixed format for each event generates easily quantifiable and tractable data sets that invite replication and facilitate easy comparison.

A second difference concerns reliance on a newspaper's annual index or
daily accounts. Most studies of race riots and civil rights’ activity used only an annual index, usually from the NYT (Spilerman 1970, McAdam 1982, Burstein 1985, Jenkins & Eckert 1986). This approach has several drawbacks. Even a quick perusal of the NYT index shows that an event can be easily confused with a “report.” For example, in practice it is hard to distinguish the NYT index entry of “Miami riots” from these alternative “events” (a) editorial comments on one or more riots, (b) subsequent stages of the same riot, (c) political speeches on a previous riot, or (d) follow-up reports such as arrests, court trials, demonstrations following arrests (Olzak & West 1987). Moreover, the actual timing of an event usually cannot be inferred from the index, which is a crucial defect when using event-history analysis. Finally, reading and coding a full account of a set of similar events is necessary for learning how or even if events are linked.

The alternative approach uses an index only as a guide to the daily accounts, suggesting candidate events and periods of unrest. Master lists from an index are sorted by date, location, and event-type, to eliminate redundancies.3 Further improvements in coding reliability occur when both the index and dailies are consulted and cross-checked by different coders (Tilly et al 1975, Franzosi 1987, Olzak & West 1987).

VARIE TIES OF EVENT ANALYSIS: RESEARCH DESIGNS

Four research designs have dominated studies of counts in collective events:

a. **Cross-sectional designs** record occurrences of events in one period for multiple units, as in Paige’s (1975) study of peasant rebellions in agricultural export sectors;

b. **Time series designs** record occurrences of events by period for a single unit, as in Snyder & Tilly’s (1972) analysis of collective violence in France from 1830 to 1930;

c. **Panel designs** (or time series/cross-section designs) record occurrences for multiple units in two or more periods, as Spilerman’s analysis (1976) of race riots among cities and over time; and

d. **Event-history designs** record the exact timing and sequence of events, as in McCarthy et al’s study (1988) of founding rates of anti-drunk driving organizations among cities.

3This is a problem since the NYT index lists events under multiple headings, but several riots can occur on one day in a city. Using the index makes it impossible to distinguish whether two reports that have the same date, location, and event type refer to the same or to different events.
A central research question in event analysis is how much to aggregate data on events over time and place. Table 1 classifies the four designs according to degree of aggregation by time and place. Cross-sectional designs aggregate completely over time. Analyses using this design commonly aggregate information on the timing of events and construct a single event-count for each locale often for a long period (Paige 1975, Jenkins 1982). In some cases the decision to aggregate counts over time reflects the absence of time series data on independent variables. Time-series designs record the total number of events over all localities or SMOs for each period (so that there is temporal aggregation within periods). Panel designs involve less aggregation because they record event counts for each unit for each period. Event-history designs involve the least aggregation: they retain exact information on timing for each locale. Readers should note that this classification into four categories is a heuristic device; these dimensions are really continua.

Analyses of data on collective events varies along another dimension: the

### TABLE 1 Classification of approaches used in event analysis

<table>
<thead>
<tr>
<th>Data Structure</th>
<th>Nonprocess Orientation</th>
<th>Process Orientation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cross section:</td>
<td>Classical regression (Paige 1975)</td>
<td>Poisson modeling (Spilerman 1970; Goldstone 1980)</td>
</tr>
<tr>
<td>complete temporal</td>
<td>Logit regressiona (Markoff 1985; McAdam 1986)</td>
<td></td>
</tr>
<tr>
<td>aggregation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>complete spatial</td>
<td></td>
<td>Poisson regression (Hannan &amp; Freeman in press)</td>
</tr>
<tr>
<td>aggregation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time series/cross section:</td>
<td>Two wave panel with lagged dependent variables used as a control (Muller 1985, Weede 1987)</td>
<td>Panel analysis used to estimate dynamic process (Spilerman 1976; Olzak 1982)</td>
</tr>
<tr>
<td>partial aggregation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event History:</td>
<td>Single point process (Olzak 1987)</td>
<td></td>
</tr>
<tr>
<td>exact timing</td>
<td>Models of transition rates: multiple units (Hannan &amp; Carroll 1981; McCarthy et al. 1988; Liebman et al. 1988)</td>
<td></td>
</tr>
</tbody>
</table>

*aLogit regression can also be considered as process oriented. See text.*
extent to which event analysis takes a process orientation. A process-oriented approach evaluates how time and covariates (including the occurrence of past events) affect the timing and sequence of repeatable events, and it attempts to explain how events unfold over time (dynamics). In contrast, other approaches, called “nonprocess” here for lack of a better label, summarize static relationships between levels or characteristics of units and event counts. Nonprocess approaches thus relate levels of unit-characteristics (aggregated with respect to time or place) to some unit-specific or period-specific event count. As in the case of the aggregation, this dimension too is a continuum.4

When past applications are considered, the process dimension can also be summarized with respect to a choice of a probability model. This choice has substantive consequences as well. Nonprocess approaches have commonly employed classical regression models which assume a normal distribution, while process models have employed discrete-state stochastic process models. Most research on events uses the former. However, data on event counts have two characteristics that make the assumption of normality suspect: event counts are non-negative and discrete. Estimation techniques that assume a normal distribution have been used more as a convenience than because they were appropriate.5 Now that software for estimating discrete-state stochastic process models is widely available, reliance on techniques based on normal distribution theory is likely to diminish.

Cross-Sectional Analysis

A variety of studies of collective action use methods that assume a normal (Gaussian) distribution to estimate the effects of community characteristics on variations in numbers of events. In cross-sectional applications this involves relating constant community characteristics to frequencies of events, aggregated over time to produce a total count for each community. This approach has been used in studies of strikes (Britt & Galle 1974, Lincoln 1978), lynchings (Inverarity 1976, Tolnay et al 1989), peasant rebellions (Chirot & Ragin 1975), organizational success (Steedly & Foley 1979), and urban protests (Eisinger 1973). Techniques used to assess the effects of covariates in these studies include correlation analysis (Jenkins & Perrow 1977), regression

4Specifically, time series and panel approaches that estimate dynamic models (models with lagged dependent variables) can be given a process interpretation (see Table 1 for examples).

5While counted data never strictly satisfy the assumptions of non-negativity and continuity, this does not make much practical difference when counts are large, as in the case of number of strikes per year in the United States. In practice, violation of this assumption of non-negativity has been circumvented by logging the counts of events or by adding 1.0 to the counts (see Snyder 1975). For most studies reviewed here this has not been done in either time series or regression analyses.
analysis (Jenkins 1982), multiple classification analysis (Corzine et al 1983), and confirmatory factor analysis (Inverarity 1976).

Studies of participation in social movements generally also use cross-sectional data. As mentioned above, analysis of surveys of only participants in efforts to learn the causes of participation produces sample-selection bias. At least five studies avoided this problem by using data on samples of individuals who did and did not participate in riots (Paige 1971, Morgan & Clark 1973, Ladner et al 1981), anti-busing protests (Useem 1980), and vandalism during riots (Berk & Aldrich 1972).

Some investigators have collapsed information on event counts into binary distinctions between “no event” and “one or more events” (Markoff 1986) or “participant” and “non-participant” (McAdam 1986). These analysts used logit regression which is more efficient than standard linear regression estimates. The gain in efficiency is important when the fraction of units with an event is small in number compared to the period of time at risk or the number of units without an event.

Used with care, cross-sectional designs maximize variability in structural properties that affect the occurrence of events. However, there are drawbacks to applying cross-sectional designs to data on events that occur over time. Specifically, this approach has largely ignored the process by which events unfold, that is, whatever affects the nature, timing, and sequence of subsequent events. A concrete research question illustrates this point. Does use of violence by the police early in a riot affect the likelihood of more violence on the part of demonstrators? Such questions direct attention to sequences or stages in an event as well as to the timing of subsequent events, and this requires models and estimation techniques with a process orientation.

But cross-sectional data can be used to estimate parameters of some simple process models. In an important example, Poisson models have been used to analyze cross-sectional data on the propensity of cities to riot (Lieberson & Silverman 1965, Spilerman 1970). Such analyses were motivated by the question asked in cross-sectional regression analyses: Do cities (or SMOs) differ with respect to some outcome? In these cases the answer was that heterogeneity affected the likelihood of an event. For instance, patterns of occurrences of race riots and organizational success were found to depart from a model of a constant rate of occurrence. Spilerman’s (1970) findings ran counter to the conventional wisdom about the relationship between poverty and the occurrence of riots (e.g. US National Advisory Commission on Civil

6As in the cases of time series and panel analysis discussed below, logit regression can be used as either a non-process or process approach. For example, Tuma & Hannan (1984: 298) point out that a logit model could be considered as an equilibrium implication in particular. However, research reviewed in this section has not considered logit regression in this way.
Disorders 1968). By using these more appropriate methods to analyze both riot and non-riot cities, he found that "racial violence is more likely where Negroes are better situated in occupational status, in education and income, and where the rate of population growth is small" (Spilerman 1970: 645).

**Single Time Series**

The time series design highlights patterns of changes over historical time. When theoretical or historical reasons suggest that structural conditions differ among periods, analysts can estimate shifts in parameters at predetermined times. Time series analysts have captured historical variation in three ways. One estimates a time trend (Snyder & Tilly 1972). A second estimates models separately by period to check whether causal structures differ by historical period (Wallace et al 1988). A third approach combines the two; it estimates an interaction between covariates and either a time trend or a set of periods (Snyder 1975).


In principle time series analysis can be used to analyze processes that underlie the flow of events. For example, Hamblin et al (1973) use time series analysis to analyze processes of diffusion or contagion. Tuma & Hannan (1984: Part III) show that time series methods can be used to analyze process models for quantitative outcomes (such as large counts). However, the research reviewed here has rarely used time series in this way. Instead much time series analysis has treated time series data very much like cross-sectional data, simply relating counts per period to levels of (or changes in levels of) covariates in each period.

However, when analysts include lagged values of the dependent variable as a regressor, the model is implicitly dynamic. That is, such models imply a time path of change in the dependent variable even if the regressors do not change over time. When analysts develop the dynamic implications of such
models—as in Hibbs' (1976) use of adaptive expectation models to explain temporal variation in strikes—time series analysis can be considered an instance of a process orientation.

Time series data can also be used to estimate parameters of discrete-state, continuous-time stochastic process models such as the Poisson. Poisson regression has been used to estimate the effects of social and economic conditions on the rate of occurrence of events using yearly counts in a single time series. For example, Hannan and Freeman (1989) used Poisson regressions (as well as event-history methods) to relate the number of labor unions to founding rates of unions. This research shows that the dynamics of social movements can be understood by focusing on (a) the internal histories of SMOs and (b) dynamics of the interdependent events that make up the cycles of social movements.

**Cross-Sectional and Time Series Analysis**

Use of both cross-sectional and longitudinal variation (panel analysis) in events represents a third approach (see Table 1). When information on events is available for two or more time periods on multiple behavioral units (cities, organizations, participants, etc), several estimation procedures can exploit more available information than does either cross-sectional or time series analysis alone. For this reason, panel analysis is particularly advantageous when analyzing rare events, such as revolutions.

Panel designs have been used frequently in cross-national analyses of political instability and collective violence. Two common methods of estimation are: (a) multiple regression analysis for two-wave panel data and (b) pooled cross-section and time series estimators that adjust for autocorrelation of disturbances. The number of time intervals constrains the choice between the two because the second approach requires data for three or more periods. Two-wave panel analysis has been used in studies of income inequality and political violence (Muller 1985, Weede 1987). Multiwave panel analyses have been used to analyze severity of race riots in cities (Spilerman 1976) and the intensity of separatist violence and demonstrations in Quebec (Olzak 1982). The latter two studies tested hypotheses that past occurrences of events as well as constant structural and demographic characteristics of cities or states affect rates of unrest.

Almost all applications of panel analysis to events and social movements include lagged values of the dependent variables as regressors. For example, studies of political violence in a period usually include as a regressor the level of political violence in some earlier period. Often the inclusion of lagged dependent variables is motivated as a “control” variable in the spirit of Lazarsfeld. The estimated effects of such variables are then rarely given a substantive interpretation, and if one is given it rarely involves process or
dynamics (such applications are therefore listed in the left column of Table 1). Other analysts begin with process models and interpret the effects of the lagged dependent variable as revealing something about the underlying dynamic process (and these are listed in the right-hand column of Table 1).

**Event-History Analysis**

Event-history analysis has been used recently to study rates of occurrence of collective actions, such as riots and protests, as well as the establishment and disbanding of SMOs. This approach involves application and generalization of methods originally designed for studying events in biostatistics and industrial engineering. This class of methods uses information on the timing and sequencing of events to estimate models of transition rates. Two general forms of event-history analysis are relevant to the study of collective action and social movements. The first involves study of recurrent events, such as riots, where the typical duration of an event is small relative to the waiting time between events. In this case the attention focuses on the rate of occurrence, that is the transition from an event count of $N$ to $N + 1$ (Coleman 1981, Amburgey & Carroll 1984, Amburgey 1986, Hannan & Freeman 1987). The second considers transitions between enduring states, such as different forms of political regimes (Hannan & Carroll 1981). The second approach can be used for studying transitions among phases in the histories of demonstrations, revolutions, and social movements (Tuma & Hannan 1984).

Event-history analysis is particularly useful for analyzing cycles or waves of protest and violence. As noted above, while the concept of cycles of social movement activity is not new, it has rarely been investigated empirically. Use of event-history analysis facilitates study of diffusion of events. Second, questions about whether contagion is restricted to particular forms or stages in collective protest can be investigated with event-history analysis as well. For example, questions about whether Civil Rights' protests spark subsequent racial conflicts, or vice-versa, can be investigated with analyses that estimate the transition rates among multiple states.

Like panel analysis, event-history analysis can use information on the timing of events in a series of localities (or SMOs) to estimate models that take both unit-specific characteristics and the timing of events. In contrast to the single point process models described above, examples of this combined cross-sectional and longitudinal design commonly analyze events in a number of cities, SMSAs, or some other subunit over a number of observation points. This approach combines the advantages of using both cross-sectional and temporal variation with the appropriateness of stochastic models that analyze events as outcomes of *processes* that unfold over time.

Event-history analysis uses more information about the transitions or sequences in collective action than do other methods. Consider the case of
causes of national coups d’État and revolutions, which have long interested social scientists (Feierabend et al. 1972). Hannan & Carroll (1981) compared panel analysis and event-history analysis in changes of national political structure. The panel analysis relates regime type (no party, military, one-party, and multiparty) at one time to regime type 25 years before in order to infer change. Event-history analysis was used to consider all transitions among the four regime-types. The two kinds of analysis come to different conclusions. So event-history analysis adds new evidence to the debate over causes of political upheaval.

Event-history methods have been applied to studies of founding and mortality rates of SMOs in communities. For example, McCarthy and associates (1988) studied founding rates of anti-drunk driving organizations in counties and learned that community characteristics such as rural/urban character and population available for recruitment affected the founding rate of these SMOs, as did grievance-driven characteristics often cited as the impetus for organizing social movements. Others have used these methods to study processes of failure or disssion in social movements. For instance, Carroll & Huo (1988) found that success of local Knights of Labor organizations in city elections increased rates of disbanding of these locals. Liebman et al. (1988) analyzed schisms in American Protestant denominations, using data on membership, organizational structure, and histories of splits within denominations. The line of research on the internal histories (including creations and disbandings) of SMOs provides new and highly flexible methods for studying various outcomes for such organizations.

FUTURE CHALLENGES TO EVENT ANALYSIS

As event analysis increasingly uses data from historical archives and newspaper records, the methods of historians and sociologists converge. This trend poses several challenges to event analysts. One challenge is to develop sociological approaches to exploiting these historical materials that are sensitive to the historical context, as Tilly (1981) notes vigorously. For example, recent studies by Tilly (1986), Markoff (1986), and Goldstone (1986) contain imaginative operationalizations and tests of historical explanations of contentious actions. The next task is to link historical arguments to existing sociological theories or to develop new theories to fit them, and to clarify historical arguments using sociological methods.

Another challenge involves improving methods for gathering data on events. An abundance of coding manuals and appendices attest to the efforts applied to improving data collection, but surprisingly few studies report attempts to use alternative sources and ways to validate the measures. If the goal of event analysis is to provide better understanding of how collective
actions unfold, then such challenges should be met. Based on the success and diversity that has characterized event analysis thus far, it makes sense to encourage progress along these frontiers as well.

ACKNOWLEDGMENTS

The author would like to thank Hans-Peter Blossfeld, Michael Hannan, Lynn Smith-Lovin, and Nancy Brandon Tuma for their helpful suggestions in preparing this paper, and Urs Schoepflin for his excellent assistance in locating studies using event analysis. This research was supported by National Science Foundation Grant SES-8420173 Rockefeller Foundation Grant (Population Sciences Division) RF 83043, and the Max-Planck-Institut für Bildungsforschung (Berlin).

Literature Cited


Franzosi, R. 1989. From words to numbers: A


COLLECTIVE EVENT ANALYSIS


