Protest event analysis and its offspring

Swen Hutter; version 01/ 2014

Chapter prepared for Donatella della Porta (ed.). *Methodological practices in social movement research.*

1   Introduction

Protest event analysis (PEA) has become a key method of social movement research over the past decades. Oliver et al. (2003) list the increasing use of PEA even among the top-four emerging trends in social movement research. The authors describe these trends as "transcending old categories and boundaries" and combining "methodological and theoretical advances" (Oliver et al. 2003, 214). The method gained ground in the 1980s and early 1990s, as Crist and McCarthy's (1996) review article on the methodological repertoires in social movement research highlights. In contrast to most other methods presented in this volume, PEA is a key methodological innovation that emerged within the social movement field itself, and has more recently been adapted and refined to study other research topics.

Researchers rely on PEA, as a type of *content analysis*, to systematically assess the amount and features of protests across various geographical areas (from the local level up to the supranational level) and over time (from short periods of time up to several decades). Usually, social movement scholars use newspaper articles as their textual sources, but the range of sources has expanded over time and covers, amongst others, police reports and information provided by new digital media. In his comprehensive introduction to content analysis, Krippendorff (2004, 18) defines content analysis as "a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use." Thus, this chapter introduces a specific *technique* and attempts to provide some practical guidelines for researchers who want to conduct a PEA.

We can certainly question the neat distinction between quantitative and qualitative content analysis since, "[u]ltimately, all reading of texts is qualitative, even when certain characteristics of a text are later converted into numbers" (Krippendorff 2004, 16). I will emphasize the interpretative work involved in any PEA, but the ultimate goal of the techniques described in this chapter is still to transform "words to numbers" (Franzosi 2004),

which then can be analyzed with the help of various statistical tools. To a certain degree, this implies a quantitative approach, but it is significant to note that PEA can be combined with various other techniques, and the data generated with its help can be combined within different research designs (on the two understandings of methods, see Wagemann in this volume).

In theoretical terms, PEA has been used largely to test and refine arguments related to the political process approach. In the words of Klandermans and Staggenborg (2002, xi f.), "Political process theory offered an innovative method: protest event analysis provided a way of measuring the effects of political opportunities in comparative designs." More specifically, researchers turned to protest event data because of its *cross-national, cross-time* and/or *cross-issue comparative character*. For example, PEA has provided answers to questions such as how national political contexts influence the levels of protest mobilization or action repertories. At the same time, the longitudinal nature of the data has allowed us to disentangle protest waves, as well as to see how protests co-vary with changes in their environment (e.g., government participation of allies, changes in the economy), or with supposed movement outcomes (e.g., decisions by parliaments, state expenditures). Moreover, PEA has been used to study how various characteristics of protest vary across issue areas.

Protest event analysis

− PEA is a type of (quantitative) content analysis

− PEA turns words into numbers

− PEA allows for the mapping of the occurrences and characteristics of protests across geographical areas, across issues/movements and over time

− PEA is closely linked to the political process approach

This chapter builds on earlier introductions to PEA—most importantly, those by Rucht and Neidhardt (1998), as well as by Koopmans and Rucht (2002). I summarize and update this work by introducing recent developments in the field, and I formulate key questions that need to be answered by those who want to conduct a PEA. Specifically, this chapter is structured as follows. The next section presents some main PEA-based research projects. The following two sections focus more closely on the 'how to do' questions: moving from data collection to data analysis. The main focus of the chapter is on aspects of data collection (i.e., on unitizing, sampling, and coding). I only briefly discuss different strategies of data analysis.

Four generations of protest event research: an overview

In this section, I highlight the wide range of questions that can be addressed by protest event data and its offspring. Furthermore, I strongly encourage anyone interested in conducting a PEA to look at the coding manuals and data of earlier projects. This helps us to see the main decisions and dilemmas of any PEA, and it may increase the comparability across datasets. Many of the existing datasets offer valuable sources for secondary analysis and the possibility of extending the data at hand.

PEA, as a form of content analysis, has several advantages: it is an unobtrusive technique, it can handle unstructured matter as data, it is context-sensitive and it can cope with large volumes of data (see Krippendorff 2004, 40ff.). To move beyond a few cases and illustrative examples is also what made PEA so attractive to social movement scholars. As Koopmans and Rucht (2002, 252) state, "PEA provides a solid ground in an area that is still often marked more by more or less informed speculation." Since early work in the 1960s and 1970s, we observe "a virtual industry of protest event data analysis" (Klandermans and Staggenborg 2002, xii). In bold strokes, one can identify four generations of PEA research (on the development of PEA research, see also Davenport 2009, 25ff; Koopmans and Rucht 2002, 232ff.; Rucht et al. 1998; Tilly 2008, 19ff.).

The first generation—"the initiators", as Rucht et al. (1998, 10) call them—consisted of researchers who were interested in various indicators for a large number of countries, or in long-term processes of social and political change. The *Handbook for Social and Political Indicators I & II* by Russett et al. (1967) as well as by Taylor and Hudson (1972) are the most prominent examples for large N-studies. Tilly and his colleagues, by contrast, were interested in the long-term trends of strike activity and political violence (Shorter and Tilly 1974; Tilly et al. 1975). However, the authors paid relatively little attention to "the selectivity of the sources, the creation of fine-grained coding categories, and the development of well-documented rules and procedures" (Koopmans and Rucht 2002, 232). This led to the first methodological debates over the selectivity of newspaper reports (see the interesting debate between Danzger 1975; Snyder and Kelly 1977).

Inspired by this research, a second generation developed, which made more extensive use of protest data. This research broke down the data according to various analytical criteria, which was possible as the categories used for the data collection were far more sophisticated. Path breaking studies were Jenkins and Perrow's (1977) work on farmers' mobilization, Kriesi et al.'s (1981) study on political activation events in Switzerland, McAdam's (1982) case study on civil rights protests in the United States, and Tarrow's (1989) study on the Italian protest cycle from 1965 to 1974. These studies focused largely on the emergence and development of social movements that were the result of 'expanding opportunities'. Furthermore, a major innovation within this generation were cross-national designs, such as the one used by Kriesi et al. (1995) in their four-country study of new social movements' mobilization. These projects focused more on the stable elements of the political context to explain differing mobilization levels and action repertoires (on environmental protest, see also Rootes 2003).

Though the second generation was sophisticated with respect to coding procedures and source selection, the authors did not invest a lot of time in qualifying the bias of their sources. Thus, a third generation assessed the bias of newspaper data more systematically. Most importantly, authors focused on the selection bias, i.e., the fact that newspapers *selectively* report on protest events, and do not provide a representative sample of all events taking place (for reviews, see Davenport 2009: 25ff.; Earl et al. 2004; Ortiz et al. 2005).[1] Furthermore, among the third generation were those who tried to be more efficient by using electronic approaches to select (and even code) protest events. Most prominent examples of half-automated procedures are (a) the *European protest and coercion data* (EPCD) collected by Francisco et al. (e.g., Francisco 1996; Nam 2006, 2007; Reising 1998, 1999), (b) Imig and Tarrow's (2001) study on European protest events, and (c) Jenkins et al.'s project for a new edition of the *Handbook for Social and Political Indicators*. All these projects are based on adapted versions of KEDS, the Kansas Event Data System software, to identify relevant protest events. Unfortunately, these projects tend to fall back on the first generation of research when it comes to the selection of sources and coding procedures and/or their value for comparative research (Imig 2001, 256f.). More modest attempts to speed up the selection process are

---

[1] Such selection bias needs to be distinguished from description and research biases. Description bias means that newspapers report false information about covered events, whereas researcher bias refers to coding and data entry errors (e.g., Franzosi 1987; McPhail and Schweingruber 1998).

simple key word searches in electronic archives (see, e.g., Maney and Oliver 2001; Strawn 2010).

Finally, there is a fourth generation that has developed since the late 1990s. Authors have moved beyond PEA by abandoning the strict focus on (aggregates of) protest events as their coding unit. On the one hand, scholars unpacked single protest events or contentious performances by focusing on action and interaction inside them (e.g., Franzosi 2004; McPhail and Schweingruber 1998; Tilly 2008). On the other hand, scholars broadened the unit of analysis beyond protest to cover a larger group of public claims making (including protest events) (e.g., Koopmans and Statham 1999, 2010a; Koopmans et al. 2005; Kriesi et al. 2012). The two approaches within the fourth generation clearly differ from each other. However, both attempt to capture the relational aspect of political contention better than traditional PEA, and their coding units share a very similar structure, i.e., subject-relation-object (see Section 3).

This short history of PEA research should emphasize (a) the broad range of questions that can be addressed with the help of this technique, (b) how the coding unit has been expanded recently, and (c) the considerable efforts devoted to evaluating the validity and reliability of the data. As Tilly (2002, 249) has emphasized in his essay on "event catalogues as theories", scholars are interested in both "a theory embodying explanation of the phenomenon under investigation, and another theory embodying explanations of the evidence concerning that phenomenon." However, in Mueller's (1997) terms, there is both a *"representational" approach* and a *"media theory" approach* to how scholars have addressed the selection bias question. The former approach accepts the selectivity of its sources but tries to hold it constant. The latter approach is more interested in precisely examining sources of media selection bias (this is what most of the cited selection bias studies do) and, eventually, incorporates these findings in general theories of protest. Davenport's (2009) recent study on the Black Panther Party is a good example of the last point. He explicitly selects sources with different bias structures to get closer to an understanding of the conflict dynamics at play. To a certain degree, members of the fourth generation of PEA research represent yet another *"public debate" approach* to the selection bias question. In this case, mass-mediated communication, or who enters the public debate in what way, is actually the key focus of the analysis.

I will come back to the selection bias discussion in Section 3, but readers who plan their own PEA could already answer the following questions: What type of research questions does the proposed project address? Which existing research comes closest? What is the approach to the selection bias usually adopted by these studies? To help newcomers in the field, Table 1 lists major projects that have information on data collection, and the data itself, published in public archives or on websites. Free and easy access to existing datasets and codebooks is very helpful since it provides important information when you plan your own project (e.g., by exemplary coding instructions). Moreover, most of these projects have not been designed to answer only a very specific research question. By contrast, these dataset are designed for secondary analysis and can be used to address a broad range of questions. In addition, some differences between the projects are not as easy to see from the published work alone, even though small changes in the data collection may lead to quite pronounced differences in the final data (for details, see Section 3). As Earl et al. (2004, 71) state in their review article, often "differences in coding criteria and procedures may account for some of what appears to be selection bias."

[Table 1 about here]

Data collection: small questions, big impacts

Let us now focus more closely on the different steps and decisions that are involved when you want to design your own study. More specifically, this section focuses on the data collection phase by breaking it down into three steps: unitizing, sampling, and coding. In this part, I want to underscore Mark Beissinger's statement,

> While certain common practices have emerged to ensure methodological rigor, the method has been operationalized differently in practically every case of its use. Standardization of categories, definitions, and approaches across objects of analysis has remained elusive, and for good reason. The advantage of the method has precisely been its adaptability to a wide variety of circumstances, depending on the researcher's purposes. […] Researchers must ultimately make decisions about which forms of action deserve to be analyzed, what features of those actions are worthy of attention, what sources should be used to gain information about these events, and how one should organize the

process of recording this information. In a well-formulated study, both theory and context must interact to inform these choices (Beissinger 2002, 460f.).

All questions related to the data collection should be taken with respect to your research question but you should not lose sight of more pragmatic considerations since PEA and its offspring are very resource-intensive techniques. This is illustrated by two examples from our research on national political change in a globalizing world (NPW) (see Hutter forthcoming; Kriesi et al. 2012). It took around five, full-time, working months to update the French protest data of Kriesi et al. (1995) for the years 1990 to 2005 (N=2,975 events)—although we relied on electronic key word searches and adopted a 'minimalist' strategy of data collection, i.e., we selected only the Monday editions of *Le Monde*. Similarly, we spent around two months of full-time work for the selection and coding of a debate on a single issue (e.g., immigration) in one country (selecting and coding approximately 300 articles from one quality newspaper; N=2,000 core sentences).

Because of the high work load of manual content analysis, I encourage all researchers to follow Krippendorff's (2004, xxii) advice, "Beginners in content analysis are advised to start with a small pilot project, to get a feel for what is involved in conducting a larger study. […] Beginning researchers will soon realize that analyzing text is not a mechanical task, and neither is designing a content analysis. Both undertakings require creativity and competence." Such a small pilot study or pre-test can save a lot of time (and other resources), especially because changing direction during a large-scale coding enterprise involves a lot of additional work—even if you simply want to change the categories used to code a single variable.

What is your main coding unit? Zooming in and out of protest events

A very crucial step is the definition of the coding units, i.e., "units that are distinguished for separate description, transcription, recording, or coding" (Krippendorff 2004, 99). To put it simply, most research in the first three PEA generations focuses on a fairly similar list of activities as coding units, which are usually labeled as 'protest events'. The list typically covers activities from the collections of signatures, to public rallies and mass demonstrations, and to more confrontational activities (e.g., blockades and occupations), as well as violent ones (e.g., physical attacks and arson). The list reflects the modern "repertoire of contention", whose development in the eighteenth and nineteenth century Tilly (1976, 1995, 2008) traced

in his path-breaking studies. Furthermore, the list resembles the standard survey questions that are used to measure "unconventional" political participation (Barnes and Kaase 1979; Teorell et al. 2007). By contrast, the fourth PEA generation either extends the type of coding units to cover a broader set of activities or it attempts to disentangle single activities covered by the traditional approach. In these cases, the coding units are no longer protest events but, for example, 'political claims', 'core sentences' or 'semantic triplets'.

Compared to the latest advances, it seems justified to argue that traditional PEA-based projects focus on a very similar coding unit. However, if we look more closely at the projects, we realize what Beissinger referred to. I illustrate this with four examples. *First*, Tilly and his collaborators' work focused on contentious gatherings, defined as "occasions in which ten or more persons outside the government gather in the same place and make a visible claim which, if realized, would affect the interests of some specific person(s) or group(s) outside their own number" (Tilly and Schweitzer 1977, 14). *Second*, the German Prodat project defines a protest event as "a collective, public action by a non-governmental actor who expresses criticism or dissent and articulates a societal or political demand" (Rucht et al. 1992, 4). *Third*, Kriesi et al. (1995) explicitly refrained from a precise definition of a protest event but used an operational approach by relying on a detailed list of specific action forms.[2] At the same time, Kriesi et al. do not specify a minimum number of participants. *Finally*, Beissinger's (2002) study on the former USSR is restricted to demonstrations (with a minimum number of 100 participants) and events of mass violence (i.e., events whose main purpose is to cause violence and which involve a minimum number of fifteen participants). As can be seen, the two key differences are (a) the action forms covered, and (b) the minimum number of participants. Tilly et al. and Beissinger clearly refer to activities where a group of people physically meet at a certain place, whereas Prodat and Kriesi et al. do not restrict their analysis to such forms but also include the collection of signatures.

---

[2] The list of Kriesi et al. is more restricted than the forms covered by Prodat, as it does not cover action forms, such as internal protest meetings, resolutions, press conferences or litigations. To be precise, part of these activities were coded as 'conventional' activities by Kriesi et al. but only for new social movement issues and, therefore, were not used in most parts of their analyses on 'unconventional activities'.

I agree with Beissinger that a key source of the differences is the differing research purpose. For example, he justifies the use of a narrow coding unit by stating, "Ideally, in this study information on other acts of contention […] should have been collected to obtain a more complete picture of how protest repertoires evolved over time. However, given the sheer number of these events and the fact that the focus of this analysis is not protest repertoires per se but rather nationalism, there were good theoretical and practical justifications for omitting them" (Beissinger 2002, 461). At the same time, I think that the differences also mirror the general problem that protest is in itself not easy to define. As Rucht et al. (1998, 9) emphasize, "unlike other forms of social and political activities, e.g., electoral behavior, protest is by its very nature a complex phenomenon." In a recent review, Opp (2009, 33ff.) again highlights the concept's ambiguity and definitional differences. To avoid conceptually imprecise concepts, Opp (2009, 38) presents a broad definition of protest, as "as joint (i.e. collective) action of individuals aimed at achieving their goal or goals by influencing decisions of a target."

In theoretical terms, such a broad definition might be beneficial. However, we need more precise guidelines for a PEA. In other words, consider carefully what type of evidence is needed to answer your research question, and what costs would be involved in extending the coding units. To illustrate this point, I calculated the differences between the four approaches based on the updated Kriesi et al. data and Prodat. Only 3.5 percent of all coded events in the Kriesi et al. data refer to the collection of signatures, but these events account for 17.6 percent of all reported participants. By contrast, demonstrations/public marches account for 47.0 percent of all events and 65.9 percent of all participants, respectively (N=19,182). Moreover, the updated Kriesi et al. data covers only 2.9 percent of events, with fewer than three reported participants, but this figure increases to 13.9 percent with less than ten participants (N=14,905 events with a reported number of participants). Similarly, 12.8 percent of all demonstrations/protest meetings, for which a number of participants was provided in the newspaper, involved not more than 100 participants (no participant figures were reported for around the same share of events). Finally, excluding the action forms not covered by Kriesi et al. from the Prodat data decreases the number of events covered in the period 1950 to 2002 by 35.1 percent (N=15,973)—strikes account for around one third of all dropped events.

Of course, whether these numerical differences really matter for your research depends on the questions you are asking and, to a large extent, on the aggregation level of the analyses (see

Section 4). For example, if we take the updated Kriesi et al. data to compare mobilization levels in the period 1990 to 2005, it makes a difference whether collections of signatures are included or not. As Figure 1 shows, the values based on all forms of action indicate some country differences that are no longer observed when excluding such activities. At the same time, if we look at the smaller sample of the Prodat data (being more comparable to the Kriesi et al. strategy), we can tell the same story about the major waves of protest in Germany (r=0.98 and 0.92 based on yearly numbers of participants and events, respectively) (see Hutter and Teune 2012). However, if we base the analysis on demonstrations/marches alone, we get similar results only for France, Austria, Britain, and Germany (r=0.96 to 0.86), whereas this is not the case for the Netherlands and Switzerland (r=0.56 and 0.39). However, it might be that you are explicitly interested in more contentious forms of action and, therefore, you actually want a view on the Swiss "protest landscape" that is not influenced by a few, very moderate collections of signatures (see, e.g., Hutter and Giugni 2009, 409).

[Figure 1 about here]

Figure 1 presents the number of participants divided by the number of inhabitants. I believe this is the best indicator for cross-national comparisons in the mobilization levels. While the number of coded protest events is a also a very good indicator with which to trace changes over time within a given country, comparing numbers of events across countries is more difficult. In the case of events, not only the size of the country accounts for differences, but the newspapers selected vary also with respect to the number of pages and articles in general, and therefore in their coverage of protests or any other events. Standardizing the number of events by the number of inhabitants, I think, is therefore a less useful strategy (but see Beissinger and Sasse 2012).

The fourth PEA generation has shifted from protest events to alternative coding units. As stated, some authors have collected data on a far broader set of coding units (including protest events), others have chosen to focus more closely on the dynamics within single events or contentious performances. Koopmans and Statham's (1999) political claim analysis (PCA), as well as Kriesi et al.'s (2012) core-sentence analysis (CSA) exemplify the first approach, whereas the work of Franzosi (2004) and Tilly (2008) exemplifies the second. The two approaches clearly differ from each other. However, both attempt to capture the relational aspect of political contention better than traditional PEA. This is reflected in the very similar

basic structure of their coding units: *subject-relation-object*. In the following, I illustrate the two approaches by briefly discussing PCA and Tilly's latest work.

Koopmans and Statham (1999) introduce PCA as a way to move beyond "protest-centric" PEA for measuring political contention, and as a way to systematically link protest events with relevant covariates. The new coding unit is an instance of claim-making (a claim) and is defined as follows: "A political claim-making act is a purposeful communicative action in the public sphere. Claim-making acts consist of public speech acts (including protest events) that articulate political demands, calls to action, proposals, or criticism, which, actually or potentially, affect the interests or integrity of the claimants or other collective actors" (Koopmans and Statham 2010b, 55). An 'ideal-typical claim' involves the following grammar sequence: "an actor, the claimant, undertakes some sort of action in the public sphere to get another actor, the addressee, to do or leaves something that affects the interests of a third actor, the object, and provides justification for why this should be done" (Koopmans and Statham 2010b, 55). By means of PCA, one can analyze the broader public debate. To give just two examples, PCA allows Koopmans et al. (2005, Chapter 5) to study the differing action repertoire of the radical right in four European countries: the share of protest events ranges from around 80 percent (Germany) to around 16 percent (France). A key finding of the Europub project listed in Table 1 is that the visibility of social movement organizations (SMOs) and other civil society actors in public debates varies across issues (they are most disadvantaged in highly Europeanized issue areas) (della Porta and Caiani 2009: Chapter 2; Koopmans 2010).

In his book *Contentious Performances*, Tilly (2008) criticizes conventional event catalogues because these do not allow us to look inside individual episodes to uncover the more fine-grained actions of, and interactions among, various actors, and because it lacks information on how various episodes are linked to each other. Tilly's main methodological suggestion, to address these shortcomings, is the coding of 'subject-verb-object' triplets (see also Franzosi 2004; Wada 2004). Empirically, he goes back to the contentious gatherings data, which he and his colleagues collected for Great Britain from 1758 to 1834. More specifically, he focuses on a variable coded for each contentious gathering (N=8,088), which reports "each distinguishable action by any formation, including the actors(s), the crucial verb, the objects(s) of the action (where applicable), and an excerpt of the text(s) […] (50,875 records)" (Tilly 2008, 36). Around 1,500 different verbs were coded and re-grouped into forty-six

aggregate categories (for example, attack, gather, request or thank). More specifically, Tilly analyzes the clustering of verbs, broad shifts in the verbs over time, as well as subject-object pairs. For example, the analysis of shifts over time highlights the rise of more modular actions at the expense of direct attacks against enemies and wrongdoers. Furthermore, the increasing importance of the Parliament in public affairs is seen as both cause and effect of the shift from direct attacks to bargaining and support (Tilly 2008, 49ff.).

Is it worth the effort? It is significant to note that both approaches tend to increase the data collection efforts. In addition, broadening the coding unit may lead to data that includes almost no protest activities and, therefore, makes the analyses of specific features of protests, as well as the co-evolution of protests and its covariates, almost impossible. For example, only 357 (or 1.7 percent) of the claims coded by the Europub project refer to protest activities (N=21,299). While this data allows us to analyze the public claims-making and the role of SMOs and NGOs within the broader public debate, it is not very useful to analyze protest activities directed towards European policies and institutions. In addition, I think that both attempts have not yet completely succeeded in carving out the relation between different protests or among protest events and other claims. Often the data analysis is restricted to aggregates of specific variables (e.g., the actors involved in a public debate) or, in Tilly's approach, the set of activities/verbs covered is too restricted to allow a more fine-grained analysis of how protest activities are embedded in the wider stream of political conflict (for an alternative unit of analysis, see Kriesi 2009).[3]

However, it is significant to note that these choices do not need to be either/or decisions. For example, in the NPW project, we were interested in the question of how relevant political parties and the electoral arena have been in articulating and mobilizing the new conflicts induced by globalization (initially, the project focused on these political actors and sites of mobilization only, see Kriesi et al. 2008). To a get a quantitative and systematic assessment

---

[3] Kriesi (2009, 347) has suggested focusing on "event quadruples" as the basic coding unit, which consist of "action1 (of claimant)-reactions (of target/public)-reactions (of claimant)-action t2 (of claimant)." By doing so, he claims that we do not lose the focus on protest events, because the action of the claimant at t1 constitutes the first protest event in the quadruplet, the action at t2 the second protest event. However, it needs to be seen how such an approach can be implemented.

for less institutionalized forms of mobilization, we decided to focus on protest events as our coding unit by extending the Kriesi et al. (1995) data. This allowed us to compare the activities within the electoral arena to what we called the protest arena (see Hutter forthcoming). However, we all know that political conflicts are not only articulated by political parties within election campaigns or by spectacular protest events. That is why we also analyzed public debates, defined as all communication related to a particular issue, irrespective of the arena in which it occurs. More specifically, we focused on three central issues related to globalization (i.e., immigration, European integration, and economic liberalization) and broadened our coding unit to so-called core sentences.[4] By doing so, we were able to identify, amongst others, the contribution of political arenas to the public debates over globalization. For example, the range of statements linked to the protest arena varied from 12.9 percent in the case of the immigration debate, via 5.6 in the economic liberalization debate to a mere 0.3 percent in the European integration debate (Helbling et al. 2012, 212).

How to delimit events in time and space. After choosing the range of events covered by your coding unit, "the delimitation of events in time and space has be to decided" (Koopmans and Rucht 2002, 236). This involves the question of the time period and the geographical area to be covered by the data, as well as the delimitation of single events from each other. For example, the NPW project asked the questions of whether and how globalization has given rise to a new cleavage in West European politics. Since we know that globalization, in its different forms, has accelerated since the late 1980s, we chose to study electoral politics and protest politics in the period from the 1970s to the mid-2000s. Since the coding of the three issue-specific debates was very time-consuming, we decided to restrict this step to the years 2004 to 2006, i.e., to a period when the new integration-demarcation cleavage is expected to have become manifest, which allows a more detailed analysis of its structure.

As Koopmans and Rucht (2002, 236) emphasize, questions of territorial delimitation need to be carefully addressed in a global age to avoid problems of methodological nationalism (for

---

[4] The core sentence approach, which has been developed by Kleinnijenhuis and his colleagues closely resembles PCA and Franzosi's approach (see, e.g., Kleinnijenhuis et al. 1997; Kleinnijenhuis and Pennings 2001). PCA and the core sentence-approach produce very similar results, but I think that the latter has some advantages because of its more operational coding unit (for an empirical comparison, see Vliegenthart et al. 2004).

some suggestion on a 'cosmopolitan political science', see Grande 2006). For example, in a recent study, we were interested the way the volume of Europeanized protests differs across countries. For this comparative analysis, we selected and protest events with 'national participants' (either reported individuals or organizations) and protests that take place on the national territory of a given state (which is the standard approach in cross-national PEA). In the case of Germany, this means that 22.6 percent of all Europeanized protests did not take place on German soil. Most of the events that did not take place in Germany but that involved German participants have taken place in Brussels or Strasbourg.

Finally, it is important to have some guidelines on how to delimit a series of events from each other. For example, The European Protest and Coercion project adopts a 24-hour rule. This means that events that last more than 24 hours are introduced as new events. To highlight the impact of such a decision, more than fifty percent of all events in this data set coded for Switzerland are the result of a single squatting event in Zurich (N=1312). In contrast to this approach, the standard solution is to code the duration of an event as a separate variable, as well as taking the timing and the locality of events as basic criteria for delimitation. For example, our updated Kriesi et al. data is based on the instruction to treat each action of a series of actions as a single event if we know that they are separated in time (different days or clearly separated periods of one day) and/or space (different cities or clearly separated parts of one city).[5] Similarly, the US dynamics of collective mobilization project's brief event guide (available on their website) states, "An event is coded as one event if (a) it includes action that is mostly continuous—no gaps of more than 24 hours in time ([…]), (b) it is located within the same city or same part of the city, and (c) it includes the same (or a subset of the same) participants whose goals are the same."

---

[5] There are two exceptions to this general rule: (a) if an article reports on several events that took place in different cities at the same time but the article contains no other specification apart from their locality, we code one protest event with the total number of participants reported (if there is however only one further information given (e.g. number of participants per event, number of injured or arrested persons), the events are treated as separate events); (b) if an article reports on different actions (chains of events), which are distinguishable from each other neither spatially nor temporally. In these cases, different actions are coded as separate protest events if there are significant changes either concerning the goals or the participants.

What are your sources? What is the selection bias of your sources?

Every PEA faces also the challenge of selecting sampling units, i.e. "units that are distinguished for selective inclusion in an analysis" (Krippendorff 2004, 98). For survey research, the sampling unit is usually the same as the coding unit. However, this is typically not the case in PEA research. Furthermore, as Krippendorff (2004, 111) aptly states, "The universe of available texts is too large to be examined as a whole, so content analysts need to limit their research to a manageable body of texts. Although attempting to answer research questions from a limited set of data introduces the specter of sampling bias, it is possible to collect data by means of sampling plans that minimize such bias." The following section focuses on the type of sources, the absolute and relative selection bias of newspaper data, as well as half-automated procedures to speed up the selection process (because often the selection of relevant articles is more time-consuming than the actual coding).

Mass media content in general, and newspapers more specifically, are still the primary source for PEA. We can select different types of newspapers. The main differences are the geographical focus and the quality press/tabloid distinction. Other mainstream media sources are international news wires (e.g., *Reuters, Agence France Press*). More recently, internet-based sources offer another source for PEA. For example, Almeida and Lichbach (2003) compare activist-based internet-sources with traditional media outlets, and find that the former report more, and a broader range of, transnational protest events than the later. In an innovative study, Earl and Kimport (2008) introduce a form of PEA that produces a generalizable sample of online protest activities. Regarding non-media sources, police archives are the most often used type of source (e.g., Fillieule 1996, 1997; Hocke 1998, 2002; McCarthy et al. 1996c; Wisler 1994; Wisler et al. 1996). In some cases, activist archives can also be helpful in collecting data on a specific movement/issue area (e.g., Foltin 2004).

Again, the number and type of sources depends on the research purpose. To be more precise, the selection of sources depends significantly on the geographical level, time period, political context, and issue area covered by a study. Let me illustrate this with our research on the new integration-demarcation cleavage. This research focuses (a) on the national level, (b) a long period from 1975 to 2005 (usually comparing five-year periods), (c) a politically stable context, and (d) all types of issues. In this case, there is no alternative to the study of national

newspapers.[6] However, as Koopmans (1995, 253) aptly states, "It is the poverty of the alternatives that makes newspapers so attractive." The major advantages of newspapers are access, selectivity, reliability, continuity over time, and ease of coding. Newspapers report on a regular basis, they are kept in public archives, and—at least in case of quality newspapers— they try to maintain their credibility by covering events accurately. Though police archives have certain advantages over newspapers (e.g., the coverage of smaller events; usually, more structured reports), they are also biased, less comparable (even within a single country), and often contain less information on certain key variables of interest (e.g., the goals of the protestors). As Myers and Schaefer Caniglia (2004, 522) state, "The police data strategy used in recent studies is not much help because it is workable only on a local level. For a national or international study, it would be impossible to locate comparable police records for the hundreds of locations involved." Similarly, international news wires might be a good source to map broad transformations, such as the rise of Europeanized protests in all EU member states. However, international news wires are not as well suited for cross-national comparisons since they often neglect protests in smaller countries (Imig 2001, 256f.).

In other contexts, you might however want to focus on multiple sources. For example, Beissinger (2002, 476) advocates the use of multiple sources in politically unstable contexts, "although scholars studying protest in advanced industrial societies prefer a single set of newspaper sources available throughout the entire period under study to ensure consistency in coverage, the reality is that in a revolutionary society like Gorbachev's USSR, this is impossible. In a revolutionary society the best strategy available to a researcher may well be a 'blanketing' strategy, utilizing multiple sources and multiple types of information whenever they are available." As stated earlier, Davenport (2009) presents the use of multiple, biased sources as a general research strategy. While this might be a way to deal with selection bias problems, I think it is most important in exactly the type of context that he studies: the peak of a highly controversial and salient conflict.

---

[6] Furthermore, the newspapers were selected with respect to six criteria: continuous publication throughout the research period, daily publication (Monday to Saturday), high quality, comparability with regard to political orientation (none is either very conservative or extremely left-wing), coverage of the entire national territory, and similar selectivity when reporting on protest events (for an empirical test of the last two criteria, see Hutter forthcoming).

Whether you use one or multiple sources, I would suggest relying on a *systematic sampling strategy* across context and over time. For example, the ECPD project by Ron Francisco, or Uba and Uggla's (2011) recent study on Europeanized protests, both rely on multiple sources from electronic archives (LEXIS-NEXIS and Factiva). Since not all sources have been electronically available over the whole research period covered, some changes in the dataset might simply be because of the number of sources and/or the restriction to certain types of languages (in the case of Uba and Uggla, to English, Spanish, and Swedish).

Any scholar who works with PEA data needs to address the *selection bias* question (no matter what kind of approach she adopts, see Section 2). In the words of Tilly (2002, 249) "anyone who builds [event catalogues] worries unavoidably about problems of selectivity, reliability, verifiability, comparability, bounding, and inclusiveness. If compilers of event catalogs do not worry about these problems, their critics surely will." Obviously, I cannot summarize the lively and controversial debate over the selection bias of newspaper data in a few paragraphs, but I do want to point to some key findings in this literature. Most importantly, note that researchers continue to disagree on how severe the selection bias is (just compare the reviews of Earl et al. 2004 and Ortiz et al. 2005). However, no researcher would claim that these events are a representative sample of all protest events that take place. The coverage is selective, but what are the main factors that predict whether an event is covered? Knowing this helps in the interpretation of PEA findings. According to Earl et al. (2004), three sets of factors predict selection bias and increase the news value of a given protest event:[7]

- *Event characteristics*: The most important characteristics that increase the likelihood of an event being covered refer to what della Porta and Diani (2006, 171ff.) call the "logic of numbers" and the "logic of damage." Many studies show

---

[7] Authors studying selection bias compare local and national newspapers (e.g., Fillieule 1997; Hocke 1998, 2002; Swank 2000), newspapers and television (McCarthy et al. 1996b), or newspapers and non-media sources (especially police archives) (e.g., Barranco and Wisler 1999; Fillieule 1996; Hocke 1998, 2002; McCarthy et al. 1996c; Oliver and Maney 2000; Oliver and Myers 1999). Recently, scholars also relied on qualitative interviews with journalists (Fillieule and Jiménez 2003). The key results of these literature are presented in Section 3.

that large and violent events are more likely to be reported than small and peaceful ones (e.g., Barranco and Wisler 1999; Fillieule 1996; Hocke 1998, 2002; McCarthy et al. 1996c; McCarthy et al. 2008; Maney and Oliver 2003; Oliver and Maney 2000; Oliver and Myers 1999). Rucht and Neidhardt (1998, 76) even state, "In the case of very large events, as in cases of violent demonstrations leading to significant damage to property and/or injuries, we can expect a total coverage even when using only one national newspaper." Other event characteristics, which increase coverage rates, refer to the presence of counterdemonstrators and police forces or sponsorship by formal organizations (e.g., Hocke 2002; McCarthy et al. 2008; Myers and Schaefer Caniglia 2004; Oliver and Maney 2000).

• *News agency characteristics*: Danzger (1975) showed years ago that the presence of a wire service in a city increases the likelihood that an event will be covered. Oliver and Myers (1999) show, for example, that 'routinized' events confirming expectations about when, how, and where events are taking place are more likely to be covered by journalists than 'non-routinized' events. Additional variables refer to audience characteristics and newspapers' self-definition. For example, local newspapers are less selective than national newspapers (e.g., Hocke 1998, 2002; Swank 2000), and liberal or extreme left newspapers are less selective than conservative papers (e.g., Eilders 2001; Koopmans 1995; Oliver and Myers 1999).

• *Issue characteristics:* Protests that resonate with more general concerns are more likely to be reported. This is what Downs (1972) calls the "issue attention cycle", and McCarthy et al. (1996a) call the "media attention cycle." In empirical research, it is difficult to identify such attention cycles outside the newspaper coverage that such cycles are supposed to influence (Ortiz et al. 2005, 401). McCarthy et al.'s (1996a) study on Washington, D.C., is most often cited as showing the effects of media attention cycles. But even though McCarthy et al. (1996a, 492) observed some effects, these effects "are dwarfed by the consequences of size on media coverage." In another local study, Oliver and Maney (2000) show that legislative conflict over an issue increases the likelihood of a protest being covered.

Overall, the results on issue characteristics are less clear-cut than on event and news agency characteristics (Ortiz et al. 2005, 401). Another crucial question is whether these biases are consistent over time. Some studies find inconsistent patterns across short periods of a week or a month (e.g., Myers and Schaefer Caniglia 2004; Oliver and Maney 2000; Swank 2000), whereas others show that the patterns of selection bias tend to be stable. This holds especially within individual newspapers, for national sources, and over longer periods of time (e.g., Barranco and Wisler 1999; McCarthy et al. 1996c; McCarthy et al. 2008). Those who find rather negative results tend to focus on the local level and cover both protest events and more 'conventional' forms of action (Oliver and Maney 2000; Oliver and Myers 1999). McCarthy et al. (2008) provided strong evidence in favor of the stability of bias. Based on data for Minsk (Belarus), the authors show that the patterns of selection bias are very stable even in a period of political transition (i.e., from 1990 to 1995).

Scholars who adopt, in particular, a 'representational' approach ask the question of whether adding more sources or sampling more articles from a specific source is really worth the investment. I believe that an answer to this question depends a great deal on the aggregation level of issues and time periods. For example, the minimalist strategy chosen by Kriesi et al. (1995) becomes more problematic when we disaggregate these variables too far. For example, tracing the development of a specific type of environmental protest over time (e.g., transportation issues), or looking at yearly changes in welfare-related protests, seems not very reasonable with this dataset. To emphasize this point, Table 2 shows that when we take a middle-range aggregation level, the minimalist strategy leads to almost the same results as the more encompassing Prodat strategy. The table presents correlation coefficients for the trends based on Prodat and the updated Kriesi et al. dataset. It is clear that the more we aggregate the time variable (moving from one-year to five-year periods), the closer the fit between the two trend lines. Based on five-year periods, the development of the absolute number of events and participants is highly correlated ($r=0.93$ and 0.95), and the salience of specific issues is even more closely related ($r>0.96$).

[Table 2 about here]

Since there is no standard solution to the selection of sources, each researcher should at least explicitly justify the selection, discuss its advantages and disadvantages, as well as refer to the literature or their empirical material to discuss how the selection of source(s) might have

affected the findings and general conclusions. However, as Earl et al. (2004, 96) state, in a historical perspective, it is "rather ironic that researchers are so concerned with selection bias". Many earlier designs were not based on systematic quantitative research, or sampled on the dependent variable (Olzak 1989, 121). Thus, the discussion of selection bias problems should also focus on relative improvements over prior research strategies. As shown the 'how bad question' depends very much on your research question and the aggregation levels of key variables.

Finally, I would like to point to the possibility of *half-automated selection strategies*. We can distinguish rather simple keyword-based searches from more advanced methods. Some scholars doubt the usefulness of keyword-based searches (e.g., Maney and Oliver 2001), but our own experiences were very positive. When we updated the Kriesi et al. (1995) data, we used a comprehensive list of keywords to be both more efficient and consistent with the manually selected data sets. For example, we performed comparability tests based on the 1993-1999 time period for Switzerland, and for two years in all the other countries. Overall, the results are good news for those relying on electronic selection since there are hardly any differences between the manual and electronic search strategies (results available upon request). However, this type of selection is still very time-consuming, since it entails looking at many false positive hits. More advanced technologies rely on text classifiers that usually work on word frequency models. First tests show that such techniques perform quite well and clearly reduce the workload involved in the selection of articles (see Wüest et al. 2013). While the half-automated selection of protest events from digital text sources works relatively well and can also be implemented quite easily in smaller research projects, the half-automated coding of events is still mainly restricted to English sources and to highly standardized types of texts (e.g., the titles of news agency reports, as used by some projects in the third generation of PEA, see Section 2).

The accuracy of the electronic archives is of course another issue that needs to be considered when you plan to use the digital versions of newspapers. For example, in the case of the United Kingdom, we had to manually select all legends to pictures since they were not systematically included in the electronic text archive of *The Guardian*. Thus, you should always check the quality of the electronic archive before the data collection. However, the quality of the digital newspaper archives has increased over time. Thus, potential differences

between printed editions and the electronic archives seem no longer such a significant issue for research focused on more recent time periods.

What specific variables do you want to code? How do you organize the coding process?

For the coding, you should prepare a codebook in which you present instructions for the delineation of events as well as for the coding of all the variables that you are interested in. This step shows clearly just how flexible PEA is, since you can gather information on a whole range of characteristics of your coding unit. For example, the final Prodat dataset covers more than 170 variables, and the Europub dataset around 120 variables. However, as stated by Koopmans and Rucht (2002, 257), we should not "just create a shopping list of items of interest" when it comes to the specific variables coded. Note that Prodat and other datasets were created precisely to answer many different research questions and, therefore, the list of variables is very long. Apart from secondary analysis, this long list can also help you to see for which variables newspapers usually provide information (for an instructive list of variables, see Rucht and Neidhardt 1998, 82).

It is important to say that coding instructions should be formulated as precisely as possible. It is always advisable to work with examples and borderline cases. Even if one person alone does the coding, clear instruction guidelines need to be written. This helps a great deal when it comes to the analyses of the data, and it makes the work more accessible and comprehensible for non-specialists. Again, the existing codebooks are very valuable sources and you should also consider making your own codebook and data available to your readers. In addition, it is important to formulate exhaustive and mutually exclusive categories for the specific variables. If multiple values for a given variable need to be coded (e.g., the goals of the protestors or the addressee), I would suggest using multiple variables and not additional categories of the same variable. It is also quite helpful for the recoding and analysis of the data when the dataset includes a string variable that contains a brief description of the event. For example, we let our coders briefly answer the following questions in this variable: Who protests? What form of action do they use? Where do they protest? What do they want? Preferably, the coder should use the words/phrases used in the article.

For the organization of the selection and coding process, it is important to treat it as a sequence of related steps. Based on their experiences with Prodat, Rucht and Neidhardt (1998,

85) present an ideal sequence of seven steps: (1) scanning and copying articles; (2) selecting articles that definitely refer to protest events; (3) sorting articles according to protest themes and campaigns; (4) reading articles over a period of several weeks or months; (5) coding protest events; (6) putting aside problematic cases for group discussion or a decision to be taken by the supervisor, (7) depositing articles in the hard copy archive. I would also recommend separating these steps. In the case of electronic searches, it is also helpful to print out longer articles, which need to be coded. Furthermore, it is clearly worth investing some time in developing a coding application: this can range from a simple Excel file with a few macros to more sophisticated programs, such as Filemaker.

It is also important to note the problem of missing information. Often, newspapers do not report on all aspects of a protest event, in which researchers are interested in. For example, in the NPW protest event data, we are missing information on the number of participants in around every fifth event coded, while the number and type of organizations involved is not reported in more than half of all the events. How missing information is treated clearly depends on the type of information (for example, missing information on the number of injured people often indicates that there were no injuries). In the NPW project, missing participation figures have been replaced by the national median of the number of participants for a given type of event (e.g., a demonstration) in that country. However, the overall number of participants is not affected as much by this decision since only a very small fraction of events is responsible for a very large number of all participants being reported as taking part in protest activities.

Intercoder and intracoder reliability is an important issue since we want to demonstrate the trustworthiness of the data generated by means of content analysis (Krippendorff 2004, 211ff.). In the end, the application of coding rules by humans will always involve subjective interpretation and thus potential sources of error, which can be minimized by clear coding instructions and regular reliability tests (both before and during the coding process). It is significant to note that reliability needs to be tested both for the identification of a relevant coding unit and for the coding of the various specific variables. Often, the consistent identification of relevant units (i.e., protest events or subject-relation-object triplets) is even more challenging than the coding of specific variables (e.g., the number of participants or form of action). To assess the reliability of a half-automated selection procedure, we gave an inexperienced coder our NPW codebook, without any further instructions, and obtained

reliability measures of about 0.70 (identification) and 0.80 (coding of a broad type of issue/action form) (based on *The Guardian*, 2010). While the reliability measure for the coding just hits the magical threshold for reliability scores, the identification was clearly below the standard baseline of 0.80. With further training and a common discussion of problematic cases, we reached values of more than 0.90 in our actual research project.

A final note on data analysis

PEA and its offspring are techniques of data collection. Since one of the key aims of PEA is to transform words into numbers, statistical tools are usually used to analyze the collected data. However, note that part of the data allows a mixed-methods approach, i.e., a combination of quantitative analyses with the presentation of more detailed qualitative material (for example, the analyses of framing strategies sometimes combines these two approaches, see Statham et al. 2010). With respect to the statistical tools used to analyze protest event data, we can essentially use the whole repertoire that empirical social research offers (for ways to represent content analysis data, see Krippendorff 2004, 191ff.). In this section, I would just like to stress the potential of quite simple descriptive analyses of PEA data or of multivariate analyses of PEA data only. Good, illustrative examples of multivariate analyses that link protest event data with covariates from other sources are Braun and Koopmans' (2010) event history analysis of instances of racist violence in German counties, as well as Walgrave and Vliegenthart's (2012) time-series analysis of the agenda-setting power of protest.

Much PEA-based research presents quite simple descriptive statistics, such as the mobilization levels shown in Figure 1. This is clearly related to the fact that PEA scholars invest a great deal in the data collection, and collect information on objects that are not as easy to grasp, or, to restate Koopmans and Rucht (2002, 252), "PEA provides a solid ground in an area that is still often marked by more or less informed speculation." Thus, simple uni- or bivariate statistics and plots often help in answering key research questions, and are an easy way to present your data. Apart from such simple graphs, we can rely on multivariate methods for analyzing the protest event data, ranging from methods that try to uncover a certain structure (e.g., factor analysis, multidimensional scaling) to those that try to test a certain structure (e.g., multiple regressions).

To give you an example, we were interested in the structure of the political space in the different political arenas. We used multidimensional scaling (MDS) techniques to uncover the structure of the partisan space and the party configuration within that space. MDS is a very flexible method, quite similar to factor analysis, and allows for a graphic representation of similarities or dissimilarities between pairs of objects (see Cox and Cox 2001; Kruskal and Wish 1984). The unfolding technique we used allows for the joint representation of actors (e.g., parties) and issues in a common space. In addition, a variant of MDS, called weighted metric multidimensional scaling, enabled us to account simultaneously for similarities between pairs of objects (party positions with respect to a set of issues in our case) and relationships (the salience of the respective issues for each party and the salience of the different parties in the party system).

[Figure 2 about here]

Figure 2 shows our results for the protest arena (it is based on party-sponsored protest events only).[8] It is shown that protest configuration is two-dimensional. The central location of cultural liberalism, welfare, and environment shows these issues are rather consensual—at least when we focus exclusively on the protest arena. The second dimension is mainly due to conflicts over immigration. As we show, not only with this graph, immigration is the only salient and contested issue *within* the protest arena, and mobilization and counter mobilization over migration-related issues open up a second dimension in the protest political space in Western Europe.

Another way to analyze protest event data is by means of multivariate regression analyses. In this case, it is, however, very important to have a clear model about what characteristic of a given protest event might influence another characteristic. A nice example is Walker et al.'s (2008) study on how action repertoires might partly depend on the institutional target that a movement selects. They explicitly discuss their ideas about what kind of other features of a protest (e.g., the protest claim and the initiating group) may influence the institutional target of a given event.

---

[8] As Rucht notes (1998, 41), "[t]hose serving as sponsors of protest almost always also participate, so that these roles can only rarely be separated when information is derived from newspapers."

With respect to party-sponsored protest events, such a strategy could be used to answer the question of whether political parties are more likely to sponsor a protest event taking place at a certain stage of the electoral cycle, when controlling for the most important other characteristics of a given event. To do so, I also performed logistic regressions and used single protest events as my cases. More specifically, the models include two independent variables related to the timing of the event: (a) has the event taken place during the election campaign or not? (b) Has the event taken place in the middle of the electoral cycle or not?[9] Regarding other characteristics of a protest event, I include information on the involvement of other formal organizations, the form of action, the number of participants, as well as the goal of the event. Table 3 shows that the hypothesis that parties' involvement in protest activities closely follows the electoral cycle is not supported. Only the French political parties tend to be more likely to support protest events that take place both during the election campaign and in the middle of the electoral cycle. In all other countries, we find only significant effects of the other event characteristics on party sponsorship. In most countries, political parties are most likely to support moderate protest events with a high number of participants, and those events that are co-sponsored by other formal organizations. Thus, it seems more the event as such that leads political parties to enter the protest arena, and not so much the relative timing of elections.

[Table 3 about here]

---

[9] Both aspects are measured with the help of dummy variables. For example, a value of 1 for the election campaign variable means that the event has taken place within the two-month period before the day of the national parliamentary Election (this is the way we defined the election campaign period in our project on the new integration-demarcation cleavage). If we observe a significant and positive effect for this variable (as in the case of France), this indicates that the likelihood of a protest event being sponsored by a political party is clearly higher shortly before Election Day than during the rest of the legislative period. Since we control for many other aspects of a protest event, we attempt to control for the effects of other factors, apart from electoral considerations, that might lead political parties to support protests.

The choice of a specific type of data analysis should not depend on a researcher's general preference. Instead, statistical techniques should be used and combined that allow the author to answer the research question and that work with the collected data. For example, multivariate regressions clearly require a minimum number of cases. While it is difficult to give precise figures, many scholars recommend at least ten times as many observations as variables in a model (for an empirical test of certain rules of thumb, see Green 1991). Similarly, the effects of "outliers", that is, extreme cases, need to be carefully examined if simple univariate measures are calculated based on a few cases only. Nonetheless, this does not mean that PEA may not be a useful strategy of data collection when dealing with "rare events"; but in that case you should search out statistical tools that are suited to deal with such situations (e.g., King and Zeng 2001) and/or combine quantitative evidence with a more qualitative and "eventful" description of your cases.

Conclusion

This chapter introduced traditional PEA and its most recent advances that either cover a broader set of coding units or try to disaggregate single protest events or contentious performances. To begin with, I presented a brief history of PEA research and introduced a few major research projects based on this technique. On the one hand, the overview should emphasize the broad range of questions that can be addressed with the help of protest event data. On the other, much of the existing datasets offer valuable sources for secondary analysis, or the possibility to extend the data in time and space. Thereafter, the main part of the chapter focused on aspects related to the data collection. More specifically, I presented the main decisions relating to the coding unit, the sampling unit, and the coding process. By doing so, I wanted to highlight that PEA is a very powerful and flexible tool for social research. However, every scholar needs to make fundamental decisions that are based on both research interests and pragmatic considerations. For example, broadening the coding unit leads to valuable information on important co-variates of protest events (e.g., elite discourse) and allows us to situate the activities of SMOs and other NGOS in the wider public debate over certain issues. At the same time, it tends to move a researcher's focus away from protest events (that are often rather rare events in these datasets).

This chapter also briefly summarized the main conclusions of the literature on the selection bias of newspaper data. As Earl et al. (2004, 77) so aptly stated, "In fact, the evidence

suggests that social movement researchers face the same question that almost all other social scientists face: Are the best available, yet imperfect, data worthy of analysis? We argue that researchers can effectively use such data and that newspaper data does not deviate markedly from accepted standards of quality […]. We conclude that researchers must approach newspaper data with a humble understanding that, although not without its flaws, it remains a useful data source. Thus, researchers should avoid both the unexamined use of newspaper data as well as blanket condemnations of its use." I believe that this advice holds not only for newspaper-based data but for protest event analysis and its offspring more generally.

Because of the resource-intensive nature of PEA and related techniques, most studies cited in this chapter came out of large-scale and often collaborative research projects. This is also reflected in the way the results are reported and published. While there are many journal articles published based on protest event data—especially related to the questions of selection bias—the most influential contributions in the field are published as single- or co-authored monographs. Publishing the findings based on PEA and its offspring in this way allowed the researchers to exploit the full descriptive potential of their data, to embed the quantitative findings in a broader theoretical argument, and to give the reader all the necessary details about the way the data were collected in the first place. However, this should not suggest that PEA cannot and should not be used in smaller research efforts. But, as suggested in this chapter, researcher should definitely test their strategy in a pretest and should invest in new and creative research designs, which may also take advantage of the already available datasets.

In general, I would like to urge social movement scholars to be creative when it comes to new coding units, to sampling strategies, as well as to the combination of different types of content analysis within a single research project. For example, we are still missing a coding unit that really links protest events within chains of various political activities in other political arenas. Moreover, it is as yet not very common to code broader public debates among various types of actors by means of PCA (or another relational type of content analysis), and then over-sample articles relating to relevant protest events and code these articles with the help of a traditional PEA. This would allow the researcher to broaden the unit of analysis without the risk of losing sight of protest politics. Furthermore, the potential of sophisticated tools for automated content analysis has not yet been fully exploited in social movement studies.

References

Almeida, Paul D. and Mark Irving Lichbach (2003). 'To the Internet, from the Internet: Comparative Media Coverage of Transnational Protests', *Mobilization* 8(3): 249-272.

Barnes, Samuel H. and Max Kaase (1979). *Political Action. Mass Participation in Five Western Democracies*. Beverly Hills: Sage.

Barranco, José and Dominique Wisler (1999). 'Validity and systematicity of newspaper data in event analysis', *European Sociological Review* 15(3): 301-322.

Beissinger, Mark (2002). *Nationalist Mobilization and the Collapse of the Soviet State*. Cambridge: Cambridge University Press.

Beissinger, Mark and Gwendolyn Sasse (2012). 'An End to Patience? The 2008 Global Financial Crisis and Political Protest in Eastern Europe. Nuffield College Working Papers in Politics, no. 2012-01.'.

Braun, Robert and Ruud Koopmans (2010). 'The Diffusion of Ethnic Violence in Germany: The Role of Social Similarity', *European Sociological Review* 26(1): 111-123.

Cox, Trevor F. and Michael A.A. Cox (2001). *Multidimensional Scaling*. London: Chapman and Hall.

Crist, John T. and John D. McCarthy (1996). '"If I Had a Hammer": The Changing Methodological Repertoire of Collective Behavior and Social Movements Research', *Mobilization* 1(1): 87-102.

Danzger, M. Herbert (1975). 'Validating Conflict Data', *American Sociological Review* 40(5): 570-584.

Davenport, Christian (2009). *Media Bias, Perspective, and State Repression*. Cambridge: Cambridge University Press.

della Porta, Donatella and Manuela Caiani (2009). *Social Movements and Europeanization*. Oxford: Oxford University Press.

della Porta, Donatella and Mario Diani (2006). *Social Movements. An Introduction*. Oxford: Blackwell Publishing.

Downs, Anthony (1972). 'Up and Down with Ecology. The "Issue-Attention Cycle"', *Public Interest* 28: 38-50.

Earl, Jennifer and Katrina Kimport (2008). 'The Targets of Online Protest', *Information, Communication & Society* 11(4): 449-472.

Earl, Jennifer, Andrew Martin, John D. McCarthy and Sarah A. Soule (2004). 'The use of newspaper data in the study of collective action', *Annual Review Of Sociology* 30: 65-80.

Eilders, Christiane (2001). 'Die Darstellung von Protesten in ausgewählten deutschen Tageszeitungen', in Dieter Rucht (eds.), *Protest in der Bundesrepublik. Strukturen und Entwicklungen*. Frankfurt: Campus Verlag, pp. 275-311.

Fillieule, Olivier (1996). *Police Records and the National Press in France. Issues in the Methodology of Data-Collection from Newspapers*. RSCAS Working Papers 1996/25. Florence: EUI.

  (1997). *Strategies de la rue: Les manifestations en France*. Paris: Presses de Sciences Po.

Fillieule, Olivier and Manuel Jiménez (2003). 'Appendix A: The Methodology of Protest Event Analysis and the Media Politics of Reporting Environmental Protest Events', in Christopher Rootes (eds.), *Environmental Protest in Western Europe*. Oxford: Oxford University Press, pp. 258-279.

Foltin, Robert (2004). *Und wir bewegen uns doch. Soziale Bewegungen in Österreich*. Wien: edition grundrisse.

Francisco, Ron (1996). 'Coercion and Protest: An Empirical Test in Two Democratic States', *American Journal of Political Science* 40(4): 1179-1204.

Franzosi, Roberto (1987). 'The Press as a Source of Socio-Historical Data: Issues in the Methodology of Data Collection from Newspapers', *Historical Methods* 20(1): 5-16.

  (2004). *From Words to Numbers. Narrative, Data, and Social Science*. Cambridge: Cambridge University Press.

Grande, Edgar (2006). 'Cosmopolitan Political Science', *British Journal of Sociology* 57(1): 87-111.

Green, Samuel B. (1991). 'How Many Subjects Does It Take To Do A Regression Analysis?', *Multivariate Behavioral Research* 26(3): 499-510.

Helbling, Marc, Dominic Hoeglinger and Bruno Wüest (2012). 'The impact of arenas in public debates over globalization', in Hanspeter Kriesi, Edgar Grande, Martin Dolezal, Marc Helbling, Dominic Hoeglinger, Swen Hutter and Bruno Wüest (eds.), *Political conflict in Western Europe*. Cambridge: Cambridge University Press, pp. 207-228.

Hocke, Peter (1998). 'Determining the Selection Bias in Local and National Newspaper Reports on Protest Events', in Dieter Rucht, Ruud Koopmans and Friedhelm Neidhardt (eds.), *Acts of Dissent. New Developments in the Study of Protest*. Berlin: Edition Sigma, pp. 131-163.

  (2002). *Massenmedien und lokaler Protest. Empirische Fallstudie zur Medienselektivität in einer westdeutschen "Bewegungshochburg"*. Wiesbaden: Westdeutscher Verlag.

Hutter, Swen (forthcoming). *Protesting Culture and Economics in Western Europe: New Cleavages in Left and Right Politics*. Minneapolis: University of Minnesota Press.

Hutter, Swen and Marco Giugni (2009). 'Protest Politics in a Changing Political Context: Switzerland, 1975-2005', *Swiss Political Science Review* 15(3): 427-461.

Hutter, Swen and Simon Teune (2012). 'Politik auf der Straße. Deutschlands Protestprofil im Wandel', *Aus Politik und Zeitgeschichte* 25/26: 9-17).

Imig, Doug (2001). 'Methodological Appendix: Building a Transnational Archive of Contentious Events', in Doug Imig and Sidney Tarrow (eds.), *Contentious Europeans: protest and politics in an emerging polity*. Lanham: Rowman & Littlefield Publishers, pp. 253-259.

Imig, Doug and Sidney Tarrow (2001). 'Mapping the Europeanization of Contention: Evidence from a Quantitative Data Analysis', in Doug Imig and Sidney Tarrow (eds.), *Contentious Europeans: protest and politics in an emerging polity*. Lanham: Rowman & Littlefield Publishers, pp. 27-49.

Jenkins, J. Craig and Charles Perrow (1977). 'Insurgency of the Powerless: Farm Workers's Movements, 1946-1972', *American Sociological Review* 42(1): 249-268.

King, Gary and Langche Zeng (2001). 'Explaining Rare Events in International Relations', *International Organization* 55(3): 693-715.

Klandermans, Bert and Suzanne Staggenborg (2002). 'Introduction', in Bert Klandermans and Suzanne Staggenborg (eds.), *Methods of Social Movement Research*. Minneapolis: University of Minnesota Press, pp. ix-xx.

Kleinnijenhuis, Jan, Jan A. De Ridder and Ewald M. Rietberg (1997). 'Reasoning in Economic Discourse. An Application of the Network Approach to the Dutch Press', in Carl W. Roberts (eds.), *Text Analysis for the Social Sciences. Methods for Drawing Statistical Inferences from Texts and Transcripts*. Mahwah: Lawrence Erlbaum Associates, pp. 191-207.

Kleinnijenhuis, Jan and Paul Pennings (2001). 'Measurement of party positions on the basis of party programmes, media coverage and voter perceptions', in Michael Laver (eds.), *Estimating the Policy Positions of Political Actors*. London: Routledge, pp. 162-182.

Koopmans, Ruud (1995). 'Appendix: The Newspaper Data', in Hanspeter Kriesi, Ruud Koopmans, Jan Willem Duyvendak and Marco Giugni (eds.), *New Social Movements in Western Europe. A Comparative Analysis*. Minneapolis: University of Minnesota Press, pp. 253-273.

(2010). 'Winners and Losers, Supporters and Opponents in Europeanized Public Debates', in Ruud Koopmans and Paul Statham (eds.), *The Making of a European Public Sphere*. Cambridge: Cambridge University Press, pp. 97-121.

Koopmans, Ruud and Dieter Rucht (2002). 'Protest Event Analysis', in Bert Klandermans and Suzanne Staggenborg (eds.), *Methods of Social Movement Research*. Minneapolis: University of Minnesota Press, pp. 231-259.

Koopmans, Ruud and Paul Statham (1999). 'Political Claims Analysis: Integrating Protest Event and Political Discourse Approaches', *Mobilization* 4(2): 203-221.

(eds.) (2010a). *The Making of a European Public Sphere*. Cambridge: Cambridge University Press.

(2010b). 'Theoretical Framework, Research Design, and Methods', in Ruud Koopmans and Paul Statham (eds.), *The Making of a European Public Sphere*. Cambridge: Cambridge University Press, pp. 34-59.

Koopmans, Ruud, Paul Statham, Marco Giugni and Florence Passy (2005). *Contested Citizenship. Immigration and Cultural Diversity in Europe*. Minneapolis: University of Minnesota Press.

Kriesi, Hanspeter (2009). 'Charles Tilly: Contentious Performances, Campaigns and Social Movements', *Swiss Political Science Review* 15(2): 341-49.

Kriesi, Hanspeter, Edgar Grande, Martin Dolezal, Marc Helbling, Dominic Hoeglinger, Swen Hutter and Bruno Wüest (2012). *Political conflict in Western Europe*. Cambridge: Cambridge University Press.

Kriesi, Hanspeter, Edgar Grande, Romain Lachat, Martin Dolezal, Simon Bornschier and Timotheos Frey (2008). *West European Politics in the Age of Globalization*. Cambridge: Cambridge University Press.

Kriesi, Hanspeter, Ruud Koopmans, Jan Willem Duyvendak and Marco Giugni (1995). *New Social Movements in Western Europe. A Comparative Analysis*. Minneapolis: University of Minnesota Press.

Kriesi, Hanspeter, René Levy, Gilbert Ganguillet and Heinz Zwicky (eds.) (1981). *Politische Aktivierung in der Schweiz. 1945-1978*. Diessenhofen: Rüegger.

Krippendorff, Klaus (2004). *Content analysis: An introduction to its methodology*. Thousand Oaks: Sage.

Kruskal, Joseph B. and Myron Wish (1984). 'Multidimensional scaling', *Sage university papers series, no. 07-011*.

Maney, Gregory M. and Pamela E. Oliver (2001). 'Finding Collective Events. Sources, Searches, Timing', *Sociological Methods & Research* 30(2): 131-169.

McAdam, Doug (1982). *Political Process and the Development of Black Insurgency 1930-1970*. Chicago: University of Chicago Press.

McCarthy, John D., Clark McPhail and Jackie Smith (1996a). 'Images of Protest: Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991', *American Sociological Review* 61(3): 478-499.

(1996b). 'Selektionskriterien in der Berichterstattung von Fernsehen und Zeitungen', *Forschungsjournal Neue Soziale Bewegungen* 9(1): 26-45.

McCarthy, John D., Jackie Smith and Mayer N. Zald (1996c). 'Accessing public, media, electoral, and governmental agendas', in Doug McAdam, John D. McCarthy and

Mayer N. Zald (eds.), *Comparative perspectives on social movements: Political opportunities, mobilizing structures, and cultural framings*. Cambridge: Cambridge University Press, pp. 291-311.

McCarthy, John D., Larissa Titarenko, Clark McPhail, Patrick S. Rafail and Boguslaw Augustyn (2008). 'Assessing Stability in the Patterns of Selection Bias in Newspaper Coverage of Protest during the Transition from Communism in Belarus', *Mobilization* 13(2): 127-146.

McPhail, Clark and David Schweingruber (1998). 'Unpacking Protest Events: A Description Bias Analysis of Media Records with Systematic Observations of Collective Action - The 1995 March for Life in Washington, D.C.', in Dieter Rucht, Ruud Koopmans and Friedhelm Neidhardt (eds.), *Acts of Dissent. New Developments in the Study of Protest*. Berlin: Edition Sigma, pp. 164-195.

Mueller, Carol (1997). 'Media Measurement Models of Protest Event Data', *Mobilization* 2(2): 165-184.

Myers, Daniel J. and Beth Schaefer Caniglia (2004). 'All the Rioting That's Fit to Print: Selection Effects in National Newspaper Coverage of Civil Disorders, 1968-1969', *American Sociological Review* 69(4): 519-543.

Nam, Taehyun (2006). 'What You Use Matters: Coding Protest Data', *PS: Political Science & Politics* April 2006: 281-287.

(2007). 'Rough days in democracies: Comparing protests in democracies', *European Journal of Political Research* 46(1): 97-120.

Oliver, Pamela E., Jorge Cadena-Roa and Kelley D. Strawn (2003). 'Emerging Trends in the Study of Protest and Social Movements', in Betty A. Dobratz, Timothy Buzzell and Lisa K. Waldner (eds.), *Political Sociology for the 21st Century*. Oxford: Jai Press Inc.

Oliver, Pamela E. and Gregory M. Maney (2000). 'Political processes and local newspaper coverage of protest events: From selection bias to triadic interactions', *American Journal of Sociology* 106(2): 463-505.

Oliver, Pamela E. and Daniel J. Myers (1999). 'How events enter the public sphere: Conflict, location, and sponsorship in local newspaper coverage of public events', *American Journal of Sociology* 105(1): 38-87.

Olzak, Susan (1989). 'Analysis of Events in the Study of Collective Action', *Annual Review of Sociology* 15: 119-141.

Opp, Karl-Dieter (2009). *Theories of Political Protest and Social Movements. A Multidisciplinary Introduction, Critique, and Synthesis*. Routledge: London.

Ortiz, David G., Daniel J. Myers, N. Eugene Walls and Maria-Elena D. Diaz (2005). 'Where do we stand with newspaper data?', *Mobilization* 10(3): 397-419.

Reising, Uwe K. H. (1998). 'Domestic and supranational political opportunities: European protest in selected countries 1980-1995', *European Integration online Papers (EIoP)* 2(5).

     (1999). 'United in Opposition? A Cross-National Time-Series Analysis of European Protest in Three Selected Countries, 1980-1995', *The Journal of Conflict Resolution* 43(3): 317-342.

Rootes, Christopher (eds.) (2003). *Environmental Protest in Western Europe*. Oxford: Oxford University Press.

Rucht, Dieter (1998). 'The Structure and Culture of Collective Protest in Germany since 1950', in David S. Meyer and Sidney Tarrow (eds.), *The Soial Movement Society. Contentious Politics for a New Century*. Lanham: Rowman & Littlefield Publishers, pp. 29-57.

Rucht, Dieter, Peter Hocke and Thomas Ohlemacher (1992). 'Dokumentation und Analyse von Protestereignissen in der Bundesrepublik Deutschland (Prodat)', *Discussion Paper III 92-103, Social Science Research Center Berlin WZB*.

Rucht, Dieter, Ruud Koopmans and Friedhelm Neidhardt (1998). 'Introduction: Protest as a Subject of Empirical Research', in Dieter Rucht, Ruud Koopmans and Friedhelm Neidhardt (eds.), *Acts of Dissent. New Developments in the Study of Protest*. Berlin: Edition Sigma, pp. 7-30.

Rucht, Dieter and Friedhelm Neidhardt (1998). 'Methodological Issues in Collecting Protest Event Data: Units of Analysis, Sources and Sampling, Coding Problems', in Dieter Rucht, Ruud Koopmans and Friedhelm Neidhardt (eds.), *Acts of Dissent. New Developments in the Study of Protest*. Berlin: Edition Sigma, pp. 65-89.

Russett, Bruce M., Hayward Alker, Karl W. Deutsch and Harold Lasswell (1967). *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Shorter, Edward and Charles Tilly (1974). *Strikes in France,1830-1968*. Cambridge: Cambridge University Press.

Snyder, David and William R. Kelly (1977). 'Conflict Intensity, Media Sensitivity and the Validity of Newspaper Data', *American Sociological Review* 42(1): 105-123.

Statham, Paul, Ruud Koopmans, Anke Tresch and Julie Firmstone (2010). 'Political Party Contestation. Emerging Euroscepticism or a Normalization of Eurocriticism?', in Ruud Koopmans and Paul Statham (eds.), *The Making of a European Public Sphere*. Cambridge: Cambridge University Press, pp. 245-273.

Strawn, Kelley D. (2010). 'Protest Records, Data Validity, and the Mexican Media: Development and Assessment of a Keyword Search Protocol', *Social Movement Studies* 9(1): 69-84.

Swank, Eric (2000). 'In Newspapers We Trust? Assessing the Credibility of News Sources that Cover Protest Campaigns', *Research in Social Movements, Conflict and Change* 22: 27-52.

Tarrow, Sidney (1989). *Democracy and Disorder: Protest and Politics in Italy 1965-1974.* Oxford: Oxford University Press.

Taylor, Charles L. and Michael C. Hudson (1972). *World Handbook of Political and Social Indicators*. New Haven: Yale University Press.

Teorell, Jan, Mariano Torcal and José Ramón Montero (2007). 'Political participation. Mapping the terrain', in Jan van Deth, José Ramón Montero and Anders Westholm (eds.), *Citizenship and Involvement in European Democracies. A comparative analysis*. London: Routledge, pp. 334-357.

Tilly, Charles (1976). 'Major Forms of Collective Action in Western Europe, 1500-1975', *Theory and Society* 3: 365-376.

(1995). *Popular Contention in Great Britain, 1758-1834*. Cambridge, MA: Harvard University Press.

(2002). 'Event Catalogs as theories', *Sociological Theory* 20(2): 248-254.

(2008). *Contentious Performances*. Cambridge: Cambridge University Press.

Tilly, Charles and R.A. Schweitzer (1977). 'Enumeration and Coding of Contentious Gatherings in Nineteenth-Century Britain. CRSO Working Paper #210, Ann Arbor.'.

Tilly, Charles, Louise Tilly and Richard Tilly (1975). *The Rebellious Century, 1830-1930*. Cambridge: Harvard University Press.

Uba, Katrin and Fredrik Uggla (2011). 'Protest Actions against the European Union, 1992–2007', *West European Politics* 34(2): 384-393.

Vliegenhart, Rens, David Hollanders, Wouter van Attefeldt, Dirk Oegema and Ruud Koopmans (2004). *"The Race is over!... But Who has Won?" An empirical comparison between three methods of content analysis: the Brent Spar case. Paper prepared for the workshop "Inhoudsanalyse (en media)"* Etmaal van de communicatiewetenschap, Enschede, The Netherlands, 18-19 November 2004.

Wada, Takeshi (2004). 'Event Analysis of Claim Making in Mexico: How Are Social Protests Transformed into Political Protests?', *Mobilization* 9(3): 241-257.

Walgrave, Stefaan and Rens Vliegenhart (2012). 'The Complex Agenda-Setting Power of Protest. Demonstrations, Media, Parliament, Government, and Legislation in Belgium, 1993-2000', *Mobilization* 17(2): 129-156.

Walker, Edward T., Andrew W. Martin and John D. McCarthy (2008). 'Confronting the State, the Corporation, and the Academy: The Influence of Institutional Targets on Social Movement Repertoire', *American Journal of Sociology* 114(1): 35-76.

Wisler, Dominique (1994). *Violence politique et mouvements sociaux: Etude sur les radicalisations sociales, 1969-1990*. Genève: Georg.

Wisler, Dominique, José Barranco, Pierre Gentile and Maya Jegen (1996). 'Etat, violence politique et interactions', *Swiss Political Science Review* 2(3): 19-45.

Wüest, Bruno, Klaus Rothenhäusler and Swen Hutter (2013). 'Using Computational Linguistics to Enhance Protest Event Analysis (January 17, 2013). Available at SSRN: http://dx.doi.org/10.2139/ssrn.2286769'

Tables and figures

Table 1: Selected list of publicly available protest event datasets

Name	Geographical scope	Time period	Coding unit	Issues covered	Sources	Homepage
Political Activation in Switzerland by Kriesi et al.	Switzerland	1945-1979	Political activation events	all issues	Several newspapers and other sources	www.unil.ch/fors
New social movements in Western Europe (NSM) by Kriesi et al.	France, Germany, Switzerland, the Netherlands	1975-1989	Protest event	all issues	One national newspaper per country	www.unil.ch/fors
National political change in a globalizing world by Kriesi et al.	NSM-update plus Austria, and the United Kingdom	1975-2005; 2004-06	Protest event; core sentence	all issues; immigration, Europe, eco. liberalization	One national newspaper per country	www.ipz.uzh.ch/ forschung/npw (PEA not yet available; currently updated to 2011)
Prodat—Protest in Germany (main project) by Rucht et al.	Germany	1950-2002	Protest event	all issues	Two national newspapers	www.wzb.eu
European Protest and Coercion data by Ron Francisco et al.	29 European countries (plus 4 Latin American countries)	1980-1995	Protest event	all issues	Reuters plus additional national newspapers	web.ku.edu/~ronfran/data
Dynamics of Collective Action by Doug McAdam et al.	United States	1960-1995	Protest event	all issues	One national newspaper	www.stanford.edu/group/ collectiveaction
Mass demonstrations and mass violent events in the Former USSR by Mark Beissinger	Former USSR	1987-1992	Mass demonstration; mass violent event	all issues	Several international and national newspapers and publications	www.princeton.edu/~mbeissin
Black Panther Party, Authority Event Catalogue by Christian Davenport	United States (Bay Area)	1967-1973	Broad definition of BPP and authority events	all issues (by and directed to BPP)	Five newspapers	www.christiandavenport.com
The Transformation of Political Mobilization and Communication in European Public Spheres by Ruud Koopmans et al.	France, Germany, Italy, the Netherlands, Spain, Switzerland, and the United Kingdom	1990, 1995, 2000-2002	Political claim	Seven issue areas	Several newspapers per country	europub.wzb.eu/

Note: For the projects in the list, the interested reader can find all information on the data collection as well as the data in public archives and/or

directly on the web. For many of the other projects that were mentioned in Section 2, it is however easy to get additional information on the data collection (and sometimes even the data) by directly contacting the authors.

Figure 1: Protest participants per million inhabitants, 1990-2005 (in thousands)
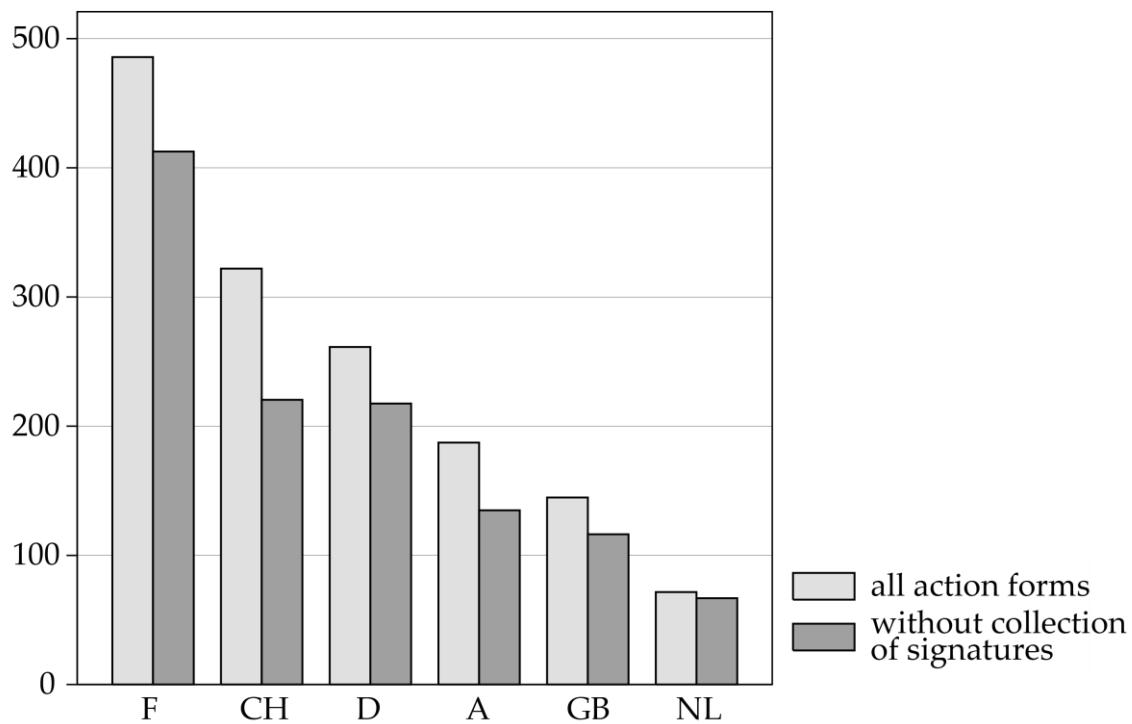
Note: The absolute number of participants is divided by the number of inhabitants in the year 2000.

Table 2: The longitudinal trends based on Prodat and Kriesi et al.

Time	N	Issues	Events	Participants
1975-97	24 years	All issues	0.85	0.79
1975-94	4 periods	All issues	0.93	0.95
1975-94	24 years	Cultural liberalism	0.89	0.94
		Environment	0.91	0.85
		Immigration	0.86	0.89
		Welfare	0.84	0.94
1975-94	4 periods	Cultural liberalism	0.99	0.99
		Environment	0.98	0.96
		Immigration	0.99	1.00
		Welfare	0.95	0.98

Note: correlation coefficients; 'events' refers to the *absolute* number of coded protest events for all issues, and the share in percent of all coded issues for the specific issues; the four time periods are 1975-79, 1980-84, 1985-89, and 1990-94.

Figure 2: Party groups in a West European protest space, 1975-2005
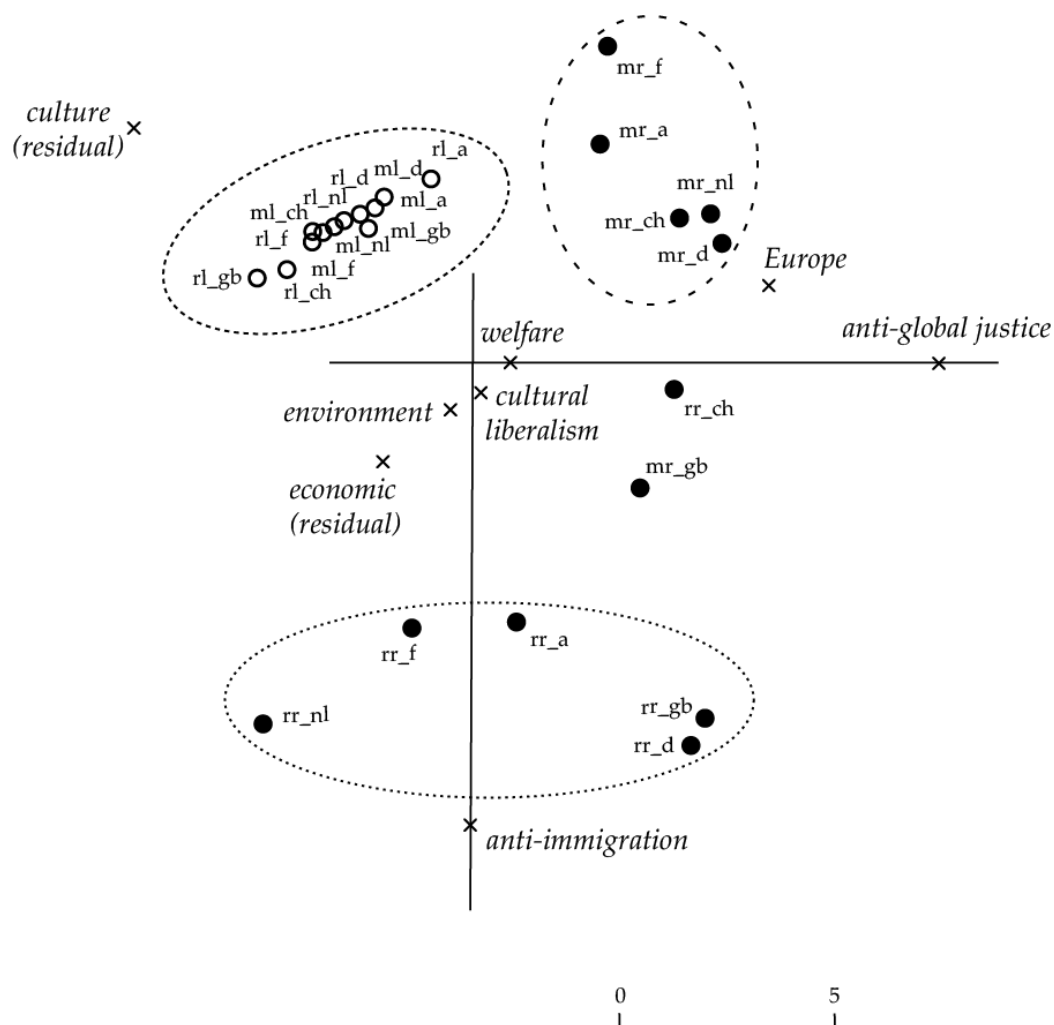
Stress-1 = 0.17

Legend: rl = radical left, ml = moderate left, mr = moderate right, rr = radial right

Note: For the MDS, party groups are weighted by their share of protest events, and the weights per country sum up to 1. The final configuration has been rotated so that the cultural conflict dimension is arranged vertically and positions supporting cultural demarcation are placed at the bottom.

Table 3: The impact of electoral cycle and event characteristics on party sponsorship, 1975–2005

	Germany	France	Netherlands	Britain
Election campaign (yes=1)	0.18	0.46*	0.07	-0.08
Middle of electoral cycle (yes=1)	0.04	0.50***	-0.07	-0.39
Event characteristics				
Supported by				
… an established interest organization	1.42***	1.73***	1.51***	0.05
… a social movement organization	0.97***	0.85***	-0.04	-0.76**
Moderate action form (yes=1)	0.64***	1.52***	0.64**	1.38***
Number of participants (1 to 5)	0.21***	0.10*	-0.05	0.41***
Issue area				
Cultural liberalism=ref.				
Immigration	-0.22	-0.19	-1.01	-1.80*
Environment	0.75***	0.74***	1.81***	2.49***
Cultural (others)	0.22	-0.47**	0.96*	0.98*
Welfare	-2.27***	-2.37***	1.08*	0.30
Economic (others)	-1.17***	-1.98***	-1.98**	1.00**
Others	-1.18***	-1.52***	-0.37	0.54
Constant	-3.88***	-4.41***	-4.01***	-5.56***
N	4,297	3,778	1,542	1,694
Pseudo-R^2	0.15	0.27	0.16	0.20

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$; logistic regression, unstandardized coefficients

Source: Updated protest event data from Kriesi et al. (1995)

Note: Logistic regression analysis (1=party-supported protest event). The number of participants has been classified into five groups: <100=1; 101-1,000=2; 1,001-5,000=3; 5,001-10,000=4; >10,000=5.