BEYOND PROTESTS: USING COMPUTATIONAL TEXT ANALYSIS TO EXPLORE A GREATER VARIETY OF SOCIAL MOVEMENT ACTIVITIES

Brayden G King and Laura K. Nelson

ABSTRACT

Social movement scholars use protest events as a way to quantify social movements and have most often used large, national newspapers to identify those events. This has introduced known and unknown biases into our measurement of social movements. We know that national newspapers tend to cover larger and more contentious events and organizations. Protest events are furthermore a small part of what social movements actually do. Without other readily available options to quantify social movements, however, big-N studies have continued to focus on protest events via a few large newspapers. With advances in digitized data and computational methods, we now no longer have to rely on large newspapers or focus only on protests to quantify important aspects of social movements. In this paper, we use the environmental movement as a case study, analyzing data from a wide range of local, regional, and national newspapers in the United States to quantify multiple facets of social movements. We argue that the incorporation of more data and new methods to quantify information in text has the potential to transform the way we both conceive of and measure social movements in three ways: (1) the type of focal social movement organization included, (2) the type of tactics and issues covered, and (3) the ability to go beyond protest events as the primary unit of analysis. In addition to demonstrating ways that the focus on counting protest events has introduced specific biases in the type of tactics, issues, and organizations covered in social movement research, we argue that computational methods can help us extract and count meaningful aspects of social movements well beyond event counts. In short, the infusion of new data and methods into

Methodological Advances in Research on Social Movements, Conflict, and Change
Research in Social Movements, Conflicts and Change, Volume 47, 13–35

ISSN: 0163-786X/doi:10.1108/S0163-786X20230000047002

social movements, peace, and conflict studies could lead us to a substantial shift in the way we quantify social movements, from protest events to everything that occurs outside of them.

Keywords: Social movements; computational text analysis; newspapers as data; protest events; tactics and strategy; environmental movement

INTRODUCTION

Social movement scholars conceive of social movements as ongoing collective action intended to create or inhibit change. When studying movements, we thus often turn our lens to visible instances of collective action, such as protests or other public events. Contentious events have long been one of the central ways social movement scholars have operationalized and measured social movements. Research on movements include studies of the antecedents and consequences of public demonstrations, boycotts, marches, and sit-ins to provide just a few examples. These events, of course, embody much of what the movement is about – signifying to the public and to the participants themselves what the movement hopes to achieve and how they hope to achieve it. Naturally, social movement scholars find these events important and worthy of study.

And yet, scholars have noted that an overemphasis on contentious events – especially on protests – gives an inaccurate picture of what social movements do, especially in contemporary society (McAdam et al., 2006). Social movements employ a variety of tactics for different strategic purposes, protest and public demonstrations being only a subset of those tactics (Nelson & King, 2020). The wide array of movement activities and movement organizations that employ them are often hidden from our scholarly view. While we know social movements are more than contentious events, how much more? What is the relative balance between participating in protests, and the rest of social movement activity? If social movement research focuses exclusively on public protest events, as scholars we fail to see (or count, if you will) less visible tactics, such as public education programs. We do not yet have a good sense of what proportion of social movement activity is captured by protest events, nor do we fully understand the quantities of the aspects of social movements we miss when we focus primarily on contentious events.

In this paper, we acknowledge the focus on contentious events is both justifiable but necessarily limiting. We seek to understand more precisely the contours of these limitations – namely, what biases do they introduce? Newspaper data have long been used to systematically count and analyze social movement activity, with a focus in particular on protests. We replicate this process using newspaper coverage of one social movement, the environmental movement, over a 25-year period (1990–2015). On a set of 187,269 news media articles from the United States from 199 different sources, we use automated event tagging to compare the nature of how the movement is covered in articles mentioning protest events, and those not mentioning protest events. Although we do not claim that these biases lead to erroneous claims about protests or other movement

events, it does lead us to make inaccurate generalizations about social movements as a broader societal phenomenon.

While assessing the nature of these biases, we also propose a methodological solution. We demonstrate how simple natural language processing and machine learning techniques, by now well developed for the social sciences, offer ways to quantify multiple dimensions of social movements beyond protest events. First, we used part of speech tagging to identify and measure the full range of tactics used by the environmental movement, contentious and noncontentious, calculating their prevalence by mention of protest events. Second, we used topic modeling to compare the issues covered across articles. And third, we used clustering techniques to identify and compare the type of organization mentioned across articles. Finally, as social movement researchers have long used large, national newspapers, such as *The New York Times*, to analyze events, scholars have also been concerned with biases arising from how these large newspapers choose what to cover and how. To situate the differences we found in articles that mentioned events and those that did not within this literature, we repeated the comparisons above for large, national newspapers compared with smaller newspapers.

In sum, in this paper we demonstrate how new methods can be used to quantify the previously unquantifiable in newspaper data. We found that only 2.4% of our articles (4,499 of 187,269) mention any sort of protest event, a stark confirmation that, even as measured by newspaper coverage – a type of media more likely to focus on public events – protests comprise a tiny minority of how social movements are publicly portrayed. Comparing the articles that mentioned protests to those that did not, we found that 63% of the environmental movement organizations (EMOs) in our data were never mentioned in an article also mentioning a protest event. Tactics, or action words, were in general more likely to be mentioned in articles that also mentioned protests, except for lifestyle and cultural tactics, which were more likely to be covered in articles not covering protests. Articles that mentioned protests were much less likely to cover conservation efforts and community-based activity and organizations, and were more likely to cover direct action, national issues, and regulations, as well as organizations that targeted institutions as opposed to community-oriented organizations. The differences we found in coverage of protest events and nonprotest coverage mirror the differences we found in coverage of the movement in large compared with smaller newspapers. The magnitude of the differences, however, was in almost every case larger when comparing protest events to nonprotest coverage than when comparing newspaper size. The focus on protests, we conclude, introduces larger biases in how we understand movement activity than the use of one or two newspapers to track protest events.

We do not think social movement scholars should abandon the study of protest events, or events more generally, as a unit of analysis. Instead, we suggest that contentious events can and should be placed into a wider context of what social movements do, how they spend their time, and how they are covered in media (beyond the large, national newspapers). Other scholars have called attention to the importance of episodes of contention (Andrews & Gaby, 2020).

We add to this that scholars should not analyze protest events separate from everything social movement organizations do between such events. New methods and data can help us do so. In the past, scholars have only been able to study these nonprotest movement activities through qualitative analysis (e.g., Polletta, 2002), but computational methods provide new ways to observe and quantify a wide variety of movement activities beyond contentious events. Thus, this paper offers tools for new ways of using quantitative analysis to study what movements do and how they do it.

CONTENTIOUS EVENTS AND SOCIAL MOVEMENT RESEARCH

The focus on contentious events has been part of social movement research at least since the resurgence of the subfield in the 1970s, when scholars began taking into account the public demonstrations of the various New Left movements that emerged during the late 1950s and 1960s (Della Porta & Rucht, 1995; Gusfield, 1970; Lipsky, 1968). Scholars conceived of social movements as one manifestation of "contentious politics," which also included other forms of political uprising such as revolutions (McAdam et al., 2001). From this tradition arose a picture of movements as contentious and reliant on protests and other extra-institutional, disruptive tactics to make their voices heard.

There is no doubt that some of the most vivid, exhilarating images we have of social movements involve contentious protests. Consider, for example, the antiwar protests by university students in the 1960s, which influenced the US government's decision to withdraw from the Vietnam War (McAdam & Su, 2002). Or, try to imagine the contemporary environmental movement without protests in which committed activists besieged forestry sites in order to prevent the logging industry from damaging the wilderness habitat (Doherty, 1999). More recently, anti-Trump activists took to the streets in the tens of thousands during the Women's March to protest the former President's sexism and misogynist policies (Fisher, 2019). The images of protestors wearing pink hats and carrying signs denouncing the president permeated the culture as a sign of political and cultural resistance.

Protests like these come to embody the grievances, claims, and emotions behind the social movements they represent. Protests serve a number of purposes. They are important tactics that social movement organizations use to mobilize participants and create solidarity. They are meant to influence the targets of the protest – that is, a legislature or a corporation – and encourage them to change a behavior, policy, or practice. Protests are also events that grab the attention of journalists, providing the social movement with a platform to broadcast their claims, more generally. For this reason, protests are among the most public and visible of tactics that social movement actors use to push forward their cause.

Given the numerous functions of protest, social movement scholars have been attracted to protest events as objects of study. Protests fit the conceptual idea that movements involve collective action, usually including conflictual or

extrainstitutional tactics, in order to promote change (Snow & Soule, 2010). Counting and studying what happens at protests is looking through a window into the soul of the movement. If we want to understand what a social movement is about, then protests are designed to tell us exactly that. As Michael Lipsky (1968, p. 1145) wrote, protest is "characterized by showmanship or display of an unconventional nature" in order to communicate with the various public audiences of a movement. Lacking power or access to influence, movements use protest to essentially subvert the system and push their ideas and goals into the realm of public debate.

But social movements do far more than protest. If we take the definition of social movements as people engaged in collective action in the pursuit of change, there are many forms of collective action that movements might take that do not involve protests or other contentious activities. If one considers the environmental movement (as we do here), the means that movements use in their cause are quite varied. The kinds of activities that movement actors might undertake include using a lawsuit to prevent toxic waste from being produced or stored near a community (EarthJustice, 2021), providing funding for expert witnesses who testify before Congress about water policy's impact on local river systems (Pitt, 2021), or educating students about what environmental justice is and how they might get involved (Chao, 2020). Some of these movement activities might be considered events, such as the filing of a lawsuit, but other activities occur over extended periods of times, such as education programs.

Social movements also build community by participating in cultural activities such as music and art festivals, they construct alternative institutions such as feminist book stores, they provide direct services to communities such as domestic violence shelters (Reger, 2012; Staggenborg & Taylor, 2005; Whittier, 2010), they raise awareness through public education, they participate in consciousness-raising and self-help activities (Taylor, 1996), and they seek to alter everyday practices and personal identities (Camp, 2005; Scott, 1985; Tugal, 2009), among many other noncontentious activities. As scholars of new social movements theorized many years ago, conflict and contention are tangential to the goals of movements that seek instead to stake out new collective identities and mobilize communities into existence (Melucci, 1980). But even social movements that are more contentious and that pursue institutional change as their primary objectives engage in a wide variety of activities. Arguably, the majority of activities done by social movement actors – namely, what actors spend most of their time doing – are not protests or events at all, but are everything that comes in between events (e.g., Polletta, 2002).

And yet often, many of these nonprotest activities do not appear in large-N datasets constructed by social movement scholars. There are several reasons for this. The first is, as discussed above, the theoretical bias in social movement research to focus on protest events as *the* engines of change. But other reasons are more practical and are related to the ways in which we collect and analyze data. First, nonprotest activities often occur behind the scenes in less visible forums. They are not designed to attract the attention of journalists. Their impact may be institutional – as, for example, when a movement organization files a lawsuit

against a major polluter – or may have a narrower scope than to influence the general public – as, for instance, when an EMO initiates an education program for teens. Second, if nonprotest activities are not as visible as protest events, they may be inherently less newsworthy. Journalists look for protests as indicators of what movements care about because they create narratives. Protests are designed to tell a story to the public. Other activities, even when they are as potentially impactful as a Congressional hearing, are less newsworthy. Finally, social movement scholars do not include them in large datasets because nonprotest activities do not fit a neat tactical category as protests do. Therefore, they are harder to systematically uncover and assess. Qualitative researchers who are deeply embedded in a particular organization or setting may be more attuned to these types of activities because they are aware of what movement organizers do on a daily basis. But when using a helicopter approach and creating a dataset of movement activities, we may miss the details of these crucial, everyday activities.

Not surprisingly, social movement scholars who study movements using quantitative analysis have heavily relied on newspapers as data sources. In the next section, we discuss newspaper data as a means to collect data on movements, note the particular biases that this archival source introduces, and discuss new methods used to overcome those biases.

Newspapers as Data

Protest events offer discrete, quantifiable units to analyze social movements, in comparative perspective and over time. Historically social movement events have been quantified via their coverage in large, national newspapers (e.g., the *Dynamics of Collective Action* [DoCA] database). Using these newspapers as data, typically *The New York Times* and/or the *Washington Post*, has noted benefits. They are a controlled source of data to reliably compare trends within and across social movements and over time. They are additionally widely read and are thus assumed to both reflect and shape national knowledge of, and opinions about, social movements. There is evidence to suggest that coverage in these large national newspapers impacts the ability for social movements, and social movement organizations, to attract resources, including participants (Amenta et al., 2009). Understanding who and what is covered in these newspapers is thus an important dynamic in social movement research.

There are also well-known biases in what gets covered in large, national newspapers, particularly around what types of events are covered in these newspapers, compared with what local newspapers cover (Myers & Caniglia, 2004). Newspapers are more likely, for example, to cover large and contentious protests than small ceremonies or speeches (Oliver & Maney, 2000). National newspapers are more likely to cover large and more violent protests compared with local newspapers (Andrews & Caren, 2010; Earl, Martin, McCarthy, & Soule, 2004). Larger social movement organizations, and larger social movements (e.g., the feminist movement or labor movement), are more likely to be mentioned in *The New York Times*, as is dramatic and disruptive activity (Amenta et al., 2009). The features of the social movement organization may also

impact which types of tactics are covered in the media. Corrigal-Brown (2016), for example, found that when Greenpeace, a confrontational EMO, called on the public to engage in environmental issues, they were more likely to be covered in large media outlets, while the reverse was true for another, nonconfrontational EMO, the World Wildlife Fund, likely because generating public action fits into the general perceptions of what Greenpeace does. Relying on a few large newspapers thus already biases how we understand social movement activity, and our reliance on newspaper data only reinforces the protest-centric approach to studying movements.

Computational Text Analysis and Its Potential for Reducing Bias in Collecting Data on Social Movements

The large, contentious events covered and counted in newspapers are, in all likelihood, a small proportion of what constitutes a social movement. But how small? And what does this mean for how we understand social movements? We propose a way to leverage new data and new methods in order to combine the strengths of using newspapers as a coherent source of comparable information about social movements, with the insights about nonprotest activity captured via qualitative methods. New sources of data and new methods to analyze that data have not only provided the opportunity to expand the sources of data used to extract information, we argue, but also to quantify and measure qualitatively different aspects of social movements than is typically done in large-N studies. More specifically, we use coverage of EMOs in a broad range of local, regional, and national newspapers in the United States to articulate a new direction for comparative social movement research, one that expands the scope of substantive information that can be extracted from newspapers as data. In doing so, we begin the process of quantifying how much, and what, is missed when the scope of social movements is defined via protest events.

Scholars have employed new data sources and new methods to mitigate the biases and limitations of counting events in large, national newspapers and as a way to diversify the way we gather data on movement activities and organizations. Digitized access to a wider array of newspapers has produced many new event databases focused on the developing world, for example, such as the Social Conflict in Africa Database and Urban Social Disturbance in Africa and Asia (Demarest & Langer, 2022). Machine learning and other natural language processing techniques have additionally enabled automating the extraction of events reported in very large collections of digitized news media, expanding the types of events and news sources for counting events. These automated and semi-automated systems use natural language processing and/or machine learning to extract events, and details about events, from news media. Most of these systems were built to detect multiple types of events, including wars or other international and domestic conflicts, terrorist events, and collective action events such as protests. Some of these systems, such as the International Crisis Early Warning System (ICEWS; O'Brien, 2010) and Global Data on Events Language and Tone (GDELT; Leetaru & Schrodt, 2013), are fully automated, using dictionaries and

other natural language processing methods to automatically extract and tag features of events. Others, such as the Social, Political, and Economic Event Database (SPEED), are semiautomated, using a combination of hand-coded events and supervised machine learning with fully automated techniques to identify events (see Wang et al., 2016 for a summary of these automated and semiautomated event detection systems).[1]

There have also been recent attempts to include social media to more comprehensively identify and count social movement protests and to leverage the power of the crowd and citizen science to capture richer and more detailed information about the events. The Crowd Counting Consortium and Count Love, for example, employ web-crawling techniques plus crowdsourced citizen scientists to identify protest activities, their location, and size in real time (for a summary of these recent advances see Fisher et al., 2019).

Focusing on the developing world, automated event detection, and crowdsourcing projects have introduced new sources of data and new methods to the social movement scholar's toolkit, circumventing the biases introduced by relying solely on large, national or international newspapers as well as the slow, cumbersome, and error-prone nature of hand-coding events (for a summary of errors introduced with hand coding see Demarest & Langer, 2022). These advances have increased the speed at which event counting can be done as well as details of events that can be quantified. While expanding the scope of events identified, these approaches still center the event, and in particular the contentious event, as the main unit of analysis for measuring social movements. As we demonstrate below, focusing on protest events perpetuates a systematic set of biases within social movement research.

THE ENVIRONMENTAL MOVEMENT

We used coverage of the environmental movement in United States newspapers as a case study of a broad and diverse social movement. The broad goal of the environmental movement is to change societal norms and governmental policies to embrace environmental sustainability, conservation of resources, public health quality, environmental justice, and environmental protection (Brulle, 2000). The environmental movement is an ideal setting in which to study the variety of ways in which news outlets cover a movement due to the variety of tactics and strategies organizations use to achieve their goals and the large number of active EMOs (Nelson & King, 2020). Similar to civil rights and women's movements, the environmental movement is also a relatively mature movement, with significant variation in organizational forms and tactics (Carmichael et al., 2012).

Like most social movements, the environmental movement meets the criteria for traditional contentious actions and collective behavior, employing large general protest actions such as the climate strikes led by Greta Thurnberg, more targeted protests such as the successful protests against the construction of the Canada and United States Keystone XL Pipeline, and direct action targeted at corporations and the public, led by well-known organizations such as

Greenpeace, People for the Ethical Treatment of Animals (PETA), and Earth First! This movement also includes noncontentious collective action, including promoting hunting and outdoor sports to fund and build conservation efforts, community-based recycling and conservation campaigns, educational activities, lifestyle tactics such as gardening and composting, and partnerships with business, including the high-profile partnership between Sierra Club and Clorox to create more eco-friendly cleaning products.

In short, the tactical diversity and general success of the environmental movement in recent decades, including consistent and broad coverage of this movement in news media (see, e.g., Amenta et al., 2009), provides an ideal case to analyze different ways in which the movement is covered in news media.

DATA

Following others who have studied social movement coverage in newspapers (e.g. Amenta et al., 2009; Corrigall-Brown, 2016), we analyzed how EMOs and their activities are covered in US newspapers. By searching for articles related to organizations rather than just to specific protests, we expand the pool of articles typically collected related to movement activity to include more nonprotest activities. We will, however, also capture a large share of the protest activity given that organizations are important sponsors of protests. We compiled a list of EMOs using two sources: a list of tax-exempt organizations related to the environment and conservation, and all organizations in the online version of the Encyclopedia of Associations produced by Gale Cengage Learning that were tagged with the keywords conservation, environment, or environmental in the subject. We found a total of 525 EMOs. While our list does not include every EMO in the United States, we believe this list is diverse, systematic, and, importantly, does not exhibit tactical or strategic sample selection bias.

We then collected news media data from two online news databases. We searched for articles that mentioned at least one EMO organization in the Nexis Uni newspaper database and the EBSCO Regional Business News database. The choice of these two databases was both practical and theoretical. They include a variety of national, regional, and local English-language newspapers and thus avoid the regional bias that is inherent when looking at a limited number of newspapers (Earl et al., 2004). As of 2016, Nexis Uni contained over 11,000 distinct sources, making it one of the most comprehensive digital databases specializing in news and business information (as well as legal information, which we did not use). To ensure adequate coverage of local and regional newspapers, we supplemented Nexis Uni with the EBSCO Regional Business News database, which includes smaller, more locally focused news and has many regional sources not present in Nexis Uni.

From these two databases, we obtained all of the articles that mentioned at least one of the 525 EMOs between 1990 and 2015.[2] We included all types of articles, including editorials, opinion pieces, and letters, as coverage of social movements, including protest events, occurs in all types of articles in the news

media.[3] The articles from Nexis Uni came from 406 distinct sources, and the EBSCO database added an additional 53 distinct sources, for a total of 459 distinct sources. To make our analysis as comparable as possible to previous social movement research relying on US newspapers (e.g., DoCA), and more recently, US newspapers and social media (e.g., *Count Love*), we restricted the sources to US news outlets. We additionally obtained newspaper circulation data on the 100 highest circulation newspapers from the Audit Bureau of Circulation, as reported on Infoplease, to compare large and small newspapers.[4] We used the year 2007 for the circulation data, falling in the middle of the years we collected. We then tagged each article as coming from a source with the 10 highest circulation numbers (hereafter Top 10), which captures the majority of national newspapers traditionally used to track events. Our complete dataset includes 187,269 articles from 199 US sources.[5] We then limited each newspaper article to only the sentences that were most directly connected to an EMO by including every sentence that mentioned at least one EMO and the subsequent nine sentences.

METHODS

Access to data and the methods used to analyze data has changed dramatically over the past decade in the social sciences. From "big data" and the computational power needed to process it (Lazer & Radford, 2017), to natural language processing and text mining (Evans & Aceves, 2016), to machine learning (Molina & Garip, 2019), research in the social sciences has developed a multitude of ways to process, analyze, and extract substantive meaning from unstructured data, such as text and images. Social movement scholars thus now have access to different ways of extracting information from data to analyze social movements. We used a combination of computational methods with qualitative coding and interpretation to measure four aspects of social movements: events, tactics, issues, and organizational type.

Protests and Other Events

In order to distinguish protests from other social movement activity, we tagged protest events in our data using the MPEDS (Machine-learning Protest Event Data System) software (Hanna, 2017a). Identifying events in newspapers, including protest events, is surprisingly difficult to do, even by hand (Demarest & Langer, 2022). The desire to be able to efficiently identify events and the details of events at scale has spurred multiple fully and semiautomatic software systems. Most existing automated event databases, such as ICEWS, GDELT, and SPEED, were created to detect all events, not just social movement events, and include international data, making them too general and broad for our purposes. MPEDS, alternatively, was created by sociologist Alex Hanna to do just one thing well: automatically identify social movement protest events from news databases such as Nexis Uni. For the purposes of MPEDS, Hanna defined a

protest event as a collective action that includes a claims-making or grievance expression that is attempting to influence social change, and one that includes an actor that is not the government or a business or other institutional leader. Protests events include things like rallies, strikes, demonstrations, picketing, and occupations (Hanna, 2017b, pp. 34–35). The software has been tested on other recent social movements such as the contemporary civil rights movement (Oliver et al., 2019) and suits our purposes in that we can use it to parse the protest events from other related social movement activity that is reported in newspapers.

To create MPEDS, Hanna and team used keyword searches to identify newspaper articles that may have been about protest activity from a wide variety of national, regional, and local newspapers in the United States. A team of trained coders then first coded whether each article included a protest event, and then for those that did, they coded details about the event (location, type, etc.). They supplemented these data with data from DoCA. Their combined hand-coded data, including DoCA articles, included 68,452 total articles, of which 2,943 (4.3%) contained a protest event. They then split these data into a training and test set, iteratively trained the MPEDS algorithm on the training set, and then calculated the *F*-score – the harmonic mean of precision and recall – to measure the accuracy of the final MPEDS algorithm. On the smallest newspaper where they had only a small training set, the *F*-score was only 0.49. On the larger sources with larger training sets, the *F*-score was between 0.59 and 0.77. Notably for our purposes, the recall was almost always higher than precision, suggesting that MPEDS more often overestimates rather than underestimates the number of protest events in a corpus. Additionally, they found some articles that contained true protest events that did not make it into the hand-coded "gold standard" DoCA dataset (what Hanna called false positives). The classification errors by MPEDS thus mimic those of human coders, and human coders are unlikely to perform substantially better than the MPEDS (or similar) algorithms.

MPEDS was trained on data very similar to the corpus used in our analysis. It is thus likely that the accuracy on our corpus is nearly identical to their reported accuracy. Nonetheless, we tested the precision on a random sample of 50 articles in our corpus identified as containing a protest event by the MPEDS algorithm, and a random sample of 50 articles identified as not containing a protest event (because so few of the articles in our corpus are about a protest event, true recall is difficult to verify). Of the 50 articles identified as containing a protest event, 36 we confirmed as containing a protest event (72%). Of the 50 articles identified as not containing a protest event, 46 we confirmed as accurate (92%). These rates are similar to the *F*-scores calculated by Hanna (2017b), and reconfirms that MPEDS is likely overestimating, not underestimating, the number of articles containing protest events in our corpus.

In addition to identifying protest events, the MPEDS algorithm also tags multiple features of the event, including issue, form (e.g., boycott, hunger strike, and demonstration), the target of the protest, the size and location of the event, and the organizations involved. For our purposes, we were only concerned with whether a news article mentioned a protest event. We thus only use the *form* tag.

Tactics

To identify the full range of tactics used by the organizations in our data, including but not limited to protests, we created a custom-made dictionary. Dictionaries, or lists of words associated with given categories, have a long history in content analysis and the social sciences (Oliver & Rahn, 2016; Schwartz & Ungar, 2015; Tausczik & Pennebaker, 2010). *Tactics* include actions such as strikes, demonstrations, petitions, and voting, as well as things like writing editorials or tweeting, but not actions such as receiving or believing. Thus, we contend that all tactics are verbs, but not all verbs are tactics. To construct our dictionary of tactics, we began by extracting all the verbs and verb phrases from the text using the standard part-of-speech tagger in Python's NLTK library. This tagger identified a total of 49,737 unique verb phrases in our data (excluding those that only occurred once, which are typically typos or mistakes).

We then went through each verb phrase by hand, classifying them as a tactic or not. This process left us with 841 unique tactics. Each author reviewed these tactics, and together we inductively identified nine tactical categories from these unique tactics: *Disruptive Protest, Non-Disruptive Protest, Political, Juridical, Verbal Statements, and Education/Raising Awareness, Business, Direct Environmental Protection, and Lifestyle/Culture* (Nelson & King, 2020). Each author then independently tagged each tactic as belonging to at least one, and up to three, of those nine tactical categories. The two authors agreed on at least one of the categories for 67% of the tactics. We took the intersection of the tagged categories for each of the tactics where the authors agreed on at least one categorization. For the remaining tactics, we mutually agreed on the best categories for each. These tagged words became our tactical category dictionary: a list of words associated with each of the nine categories.[6] To confirm the validity of these categories, we followed this analysis with a qualitative reading of texts with the most frequent mentions of words from each tactical category.

Issues

To identify the general issues – or topics – addressed in the news articles in our data, we used topic modeling. Topic models are by now a standard tool used to identify and summarize themes in large corpora (Bonilla & Grimmer, 2013; Grimmer & Stewart, 2013; Mohr & Bogdanov, 2013; Roberts et al., 2014). Topic models take as input the number of topics to produce, determined by the researcher, and uses that number to output (1) a weighted distribution of terms for each set topic, suggesting themes in a corpus, and (2) a weighted distribution of topics for each document, suggesting what themes each document covers. We used a variation of probabilistic topic modeling called *structural topic modeling* (STM) (Roberts et al., 2014). STM allows the inclusion of covariates when estimating topics. As language changes over time, to improve our topics, we included publication year as the only covariate. As we did not know the exact number of topics that would produce a coherent summary of the corpus, we took a data-driven approach. We first produced a 200-topic model, and then used hierarchical clustering to further reduce these 200 topics into a coherent thematic

space. To produce these thematic topic clusters, each author independently examined the hierarchical clustering of the 200-topic model, and using qualitative analyses of the words associated with each topic as well as the visual cutoff method often used in clustering analysis, independently decided the thirteen-cluster topic solution produced the most semantically coherent and comprehensive themes. Each document then was represented as a weighted distribution across these 13 topic clusters, allowing us to compare broad themes across sets of documents.

By examining top weighted words for each topic cluster, and reading top weighted documents for each cluster, we labeled the 13 topic clusters with substantively descriptive labels: *Business Sustainability*, *Community Organizing*, *Conservation*, *Individual Action*, *Community Sustainability*, *Species Preservation*, *Local Regulation*, *Local Conservation*, *National Politics*, *Direct Action*, *Public Interest*, and *Regulation*. To validate the clusters and the labels we gave the clusters, we followed DiMaggio et al. (2013) and compared the prevalence of each cluster across two well-documented EMOs: Greenpeace and Sierra Club. These two organizations are prevalent in our corpus and have very different strategies. We know Greenpeace does more direct action and tactics aimed at increasing public awareness of issues, while Sierra Club focuses more on conservation and works on local issues. If our topic clusters are identifying coherent themes in our corpus, the prevalence of the *Direct Action* and *Conservation* topic clusters should be substantially different in articles that mention Greenpeace and those that mention Sierra Club. This is indeed what we found, suggesting our clusters have substantive validity. The *Direct Action* cluster comprised 13% of the words in the articles that mentioned Greenpeace compared to only 2% of those that mentioned Sierra Club, and the *Conservation* topic comprised 13% of the words in the articles that mentioned the Sierra Club and only 6% of those that mentioned Greenpeace.

With our data-driven clustering of topics, itself a qualitative validation of the model, and the comparison of Greenpeace and Sierra Club, we assume that, while these clusters do not represent the only way to thematically cluster the text, the topic clusters are capturing real discursive themes in the corpus. We used the document by topic cluster weight to assess whether there are systematic thematic differences between the articles covering the environmental movement by mentions of protest events and by newspaper type.

Organizational Type

We were additionally interested in whether organizations that were mentioned alongside similar issues (as measured by the 13 topic clusters) could be clustered into an organizational typology. To explore this option, we concatenated all documents that mentioned an organization together into an organizational document, one for each organization in our data, and calculated the distribution of each constructed document over our 13 topic clusters. The result was an EMO by 13-dimensional matrix, with the cells consisting of the corresponding topic weights for each organization's extended document. We then used Pearson's

correlation coefficient to calculate whether two organizations' discursive vectors were correlated, producing a similarity measure between each pair of organizations. We assigned a negative (−1) or positive (1) correlation between two organizations if the correlation coefficient was significant at the $p < 0.05$ level, in the respective direction, and no correlation (0) if the p-value was greater than the 0.05 cutoff. Using this organization-by-organization similarity matrix, we hierarchically clustered organizations using the nearest point algorithm and Euclidean distance (Gordon, 1996).

We followed this clustering analysis with a qualitative analysis of the mission statements from the most frequently mentioned organizations in each cluster. We first read the mission statements and the websites from the 10 most frequently mentioned organizations in each high-level cluster, and then did the same for the next three levels into the hierarchical cluster. As we did so, we wrote brief descriptions of any patterns we found connecting the most frequent organizations in each cluster. After completing this exercise for all three levels, we then reviewed our descriptions to identify patterns within and across the different clusters.

We found three substantially distinct, meaningful clusters of organizations in our data, corresponding to what level of society the organization viewed as responsible for change. We labeled these three clusters: *Community Mobilization*, capturing organizations that saw the community as responsible for change, *Personal Transformation*, for those who believed individuals were the main change actors, and *Institutional Change*, capturing those that believed institutional change – for example, changing governmental policies or the structure of corporations – as the only path forward for lasting change (see Nelson & King, 2020 for more on these clusters).

RESULTS

Protest Events

Even given the fact that newspapers are more likely to cover protest events compared to other movement activities, we found that coverage of protest events comprised a miniscule proportion of coverage of the movement as a whole. Only 2.4% of our articles, 4,499 of the 187,269, mentioned any kind of social movement protest event, as identified using the MPEDS algorithm.

Table 1 shows the count of articles mentioning a protest event by event type (the *form* tag identified via the MPEDS algorithm). The majority of the protests identified were rallies or demonstrations, followed by blockades or other forms of disruptions, strikes or walkouts, marches, and symbolic actions. Top 10 newspapers were slightly more (but statistically significantly) likely to cover protest events, with around 3% of the articles in the Top 10 newspapers mentioning protests compared to around 2% in smaller newspapers (see Fig. 1).

Of the 525 EMOs we identified, 455 were mentioned in at least one article in our US news media data. Of these, only 168, or 37%, were mentioned in an article that also mentioned a protest event, and 358, or 79%, were mentioned at least once in the Top 10 newspapers. By numbers alone, focusing on protest events

Table 1. Count of Articles Mentioning Events by Protest Event Type.

Rally/demonstration	3,702
Blockade/slowdown/disruption	258
Strike/walkout/lockout	201
March	189
Symbolic display/symbolic action	107
Boycott	27
Occupation/sit-in	7
Riot	4
Hunger strike	4

only would have missed a large proportion of articles covering the environmental movement during this period and would have missed a majority of organizations involved in the movement.

Tactics

We found differences in the types of tactics mentioned in articles that also mentioned a protest event compared with articles that did not mention a protest

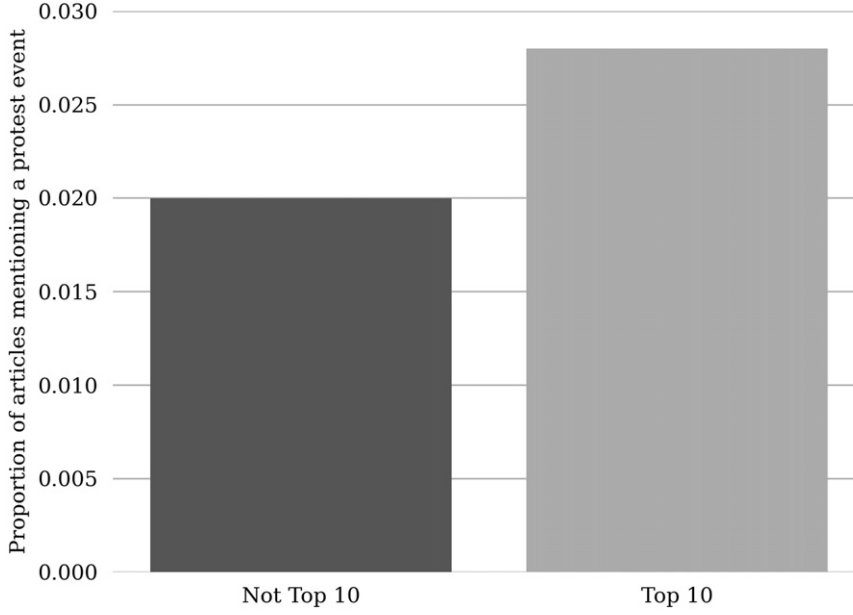

Fig. 1. Proportion of Articles That Mention a Protest Event by Newspaper Size. *Note:* This figure shows the average proportion of articles that mention protest events in top 10 newspapers and not top 10 newspapers (measured by US circulation). Events were automatically detected using MPEDS (Hanna, 2017a). Only 2.4% of all articles in the corpus mention a protest event.

event. Articles that mentioned protests were more likely to mention nondisruptive protest tactics (6.2 words per 1,000 words in articles that mention events compared with 3.5 in nonprotest articles), disruptive protest tactics (11.6 in articles that mention events compared with 3.9), and juridical tactics (7.7 compared with 5.1) (see Fig. 2). Articles that did not mention events were more likely to mention lifestyle and cultural tactics (5.8 words per 1,000 compared with 4.5 in articles that mentioned events).

We found similar differences in the articles from the Top 10 newspapers compared with the rest. By every measure, however, the differences in coverage of protest events and nonprotest activity were much larger compared with the differences in newspaper size (see Fig. 2). Articles in Top 10 newspapers were slightly more likely to mention disruptive protest tactics (4.6 words per 1,000 words in Top 10 newspapers compared with 4.0 in other newspapers) and those not in the Top 10 slightly more likely to mention lifestyle and culture tactics (5.9 compared with 5.2 in Top 10 newspapers).

Issues

The issues covered in articles about protest events also differed from those that covered nonprotest activity and Top 10 and non-Top 10 newspapers. Fig. 3 shows the average proportion of words in each article associated with each of the 13 topic clusters. Articles that mentioned protest events were much more likely to cover direct action and were more likely to cover regulation, public interest topics, and individual actions. Articles that did not mention protest events were much more likely to cover conservation, and more likely to cover community organizing, the preservation of species, local regulation and conservation, and community and business sustainability. Like articles that mentioned events, articles in the Top 10 newspapers were more likely to cover regulation, public

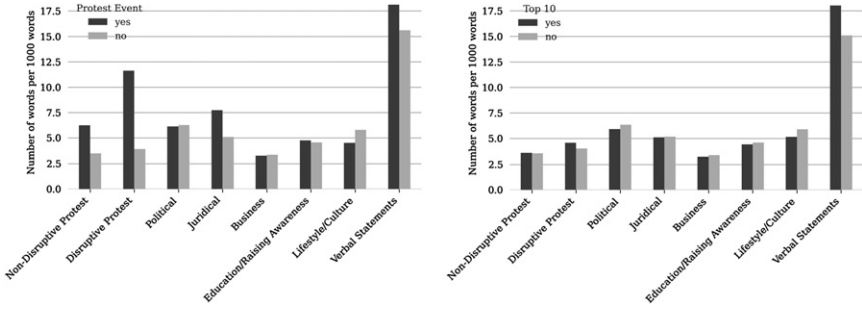

Fig. 2. Number of Words per 1,000, by Tactical Category and Mention of Protest Events (Right) and Newspaper Type (Left). *Note*: This figure shows the number of words per 1,000 words in each tactical category in all articles that mentioned protest events and those that did not (left) and all articles from the top 10 newspapers (measured by US circulation) and those not in the top 10 (right). Tactical categories were created by hand-classifying verbs into tactical categories.

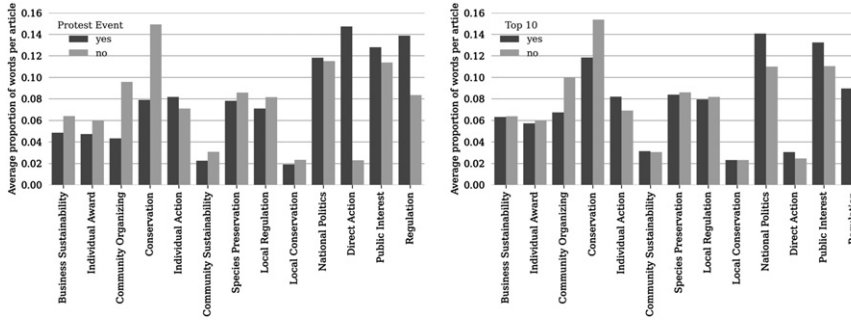

Fig. 3. Average Proportion of Words per Article Associated With 13 Topic Clusters, by Mention of Protest Events (Left) and Newspaper Type (Right). *Note:* This figure shows the average proportion of words associated with each topic cluster per article in articles that mentioned protest events and those that did not (left) and all articles from the top 10 newspapers (measured by US circulation) and those not in the top 10 (right). Topic clusters were produced by choosing a cutoff from a hierarchical cluster of a 200-topic Structural Topic Model produced on the entire corpus. The topics were labeled by the authors and are meant to suggest the content of each topic.

interest, and individual actions. Top 10 newspapers were also more likely to cover national politics. Smaller newspapers were more likely to cover environmental issues more pertinent to local communities, such as conservation.

Organizational Type

There were differences in the type of organization covered across article and newspaper type as well. Echoing the results from the topic cluster comparison, articles that mentioned protest events were 1.5 times more likely to cover organizations in the *Institutional Change* cluster compared with articles that did not mention protest events, while articles that did not mention protests were 2.4 times more likely to mention organizations in the *Community Mobilization* cluster compared with those that mentioned an event (see Fig. 4, left side). One reason for this difference may be due to the reasons that movement organizations engage in protests: they protest when trying to grab the attention of institutions, such as governments and corporations. Community organizing and conservation efforts, by contrast, seem to involve local efforts that do not involve highly visible protest tactics.

This pattern was the same when comparing newspaper size (Fig. 4, right side), but, similar to coverage of tactics and issues, the magnitude was much smaller. Newspapers in the Top 10 were 1.2 times more likely to mention organizations in the *Institutional Change* category and 1.2 times less likely to mention organizations in the *Community Mobilization* category.

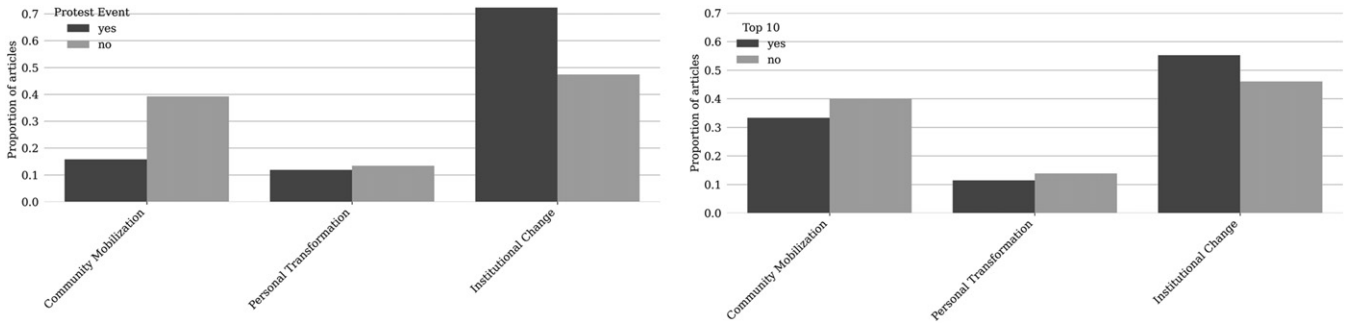

Fig. 4. Proportion of Articles Mentioning Organizational Types, by Mention of Protest Events (Left) and Newspaper Type (Right). *Note:* This figure shows the proportion of articles that mentioned each organizational type in articles that mentioned protest events and those that did not (left) and all articles from the top 10 newspapers (measured by US circulation) and those not in the top 10 (right). Organizational type was determined by clustering the organization based on their similarity in coverage across thirteen topic clusters, produced by clustering a 200-topic Structural Topic Model. The cluster names were determined by the authors by analyzing organizational mission statements from the most frequently mentioned organization in each cluster.

DISCUSSION AND CONCLUSION

Protest events are a crucial aspect of social movements; yet, protests are only a small proportion of what social movements do on a day-to-day basis. Our goal was to use coverage of the environmental movement in US newspapers as data – common in social movement research – to better understand the precise contours of the limitations of focusing primarily on protest and other contentious events. Using automated event detection software, what we found was stark: only a very small proportion (2.4%) of articles mentioned any protest event, and the majority of EMOs in our data (63%) were never mentioned in an article that mentioned any protest event. By numbers alone, we potentially miss the vast majority of media coverage and information about the majority of organizations when the focus is primarily on protest events.

Substantively, we found systematic differences in tactics mentioned, issues addressed, and organizational type covered in articles that mentioned protests compared to those that did not. The differences across all three of these ways of measuring social movements pointed to a common theme: articles that mentioned protest events were less likely to cover community-based organizations, community action, conservation efforts, and business sustainability, while overemphasizing national politics, regulation, public interest issues, and organizations that target institutions rather than the community.

When we imagine social movements, we often imagine the type of movement captured in news articles mentioning protests. Movements that target political and economic structures, governments and corporations, and national public issues, and those that do so using public protests and often illegal and contentious actions. By selecting to measure movements only through the observation of protest events, scholars reinforce and uphold this view of social movements.

While scholars, particularly qualitative scholars, have long identified and analyzed many other aspects of social movements aside from their participation in protest events, the ability to capture a variety of movement activities in large-N datasets has been more difficult. And yet, our findings suggest that the relative balance of participation in protest events compared with nonprotest activity is skewed heavily toward nonprotest activity. While we should not abandon analyzing protests, protest events happen within and around the context of these community-based organizations, actions, and issues. This aspect of social movements deserves as much, if not more, attention from social movement scholars in the future.

A major premise of this paper is that different data sources and new methods open up the possibility to quantify a multitude of aspects of social movements, beyond protests. Most large-N protest studies rely on nationally circulated newspapers with a large readership. As we point out in our study, these types of media are more likely to cover protest events than other movement activities. Regional and local newspapers by comparison are more likely to cover community-based organizations, issues, and actions, and these newspapers are increasingly available in digitized format and over longer time periods, enabling the type of longitudinal and geographic comparisons of nonprotest activity that

has long defined our understanding of protest events. Social media may potentially provide an even broader view of social movements, though we do not yet know whether and what biases might be introduced via social media data. Future research could explore social media as data, including the social media presence of social movement organizations themselves, to compare against newspaper coverage of movements.

Of course, the problem introduced by trying to incorporate a variety of local news sources is that the sheer number and variety of text may be overwhelming even for a large research team. New methods developed by computer and information scientists provide the opportunity to quantify meaningful aspects of social movements in these newly available data. We demonstrated how three types of computational text analysis methods can be used to extract different types of information from text and the substantive meaning attached: part of speech tagging to identify and count tactics, topic modeling to quantify issues or themes in text, and clustering to identify and compare organizational clusters within a social movement.

There are, of course, many different methods available to social scientists that can be used to quantify other aspects of social movements. For example, named entity recognition could be used to identify the target of social movements, phrase mining can identify important concepts or ideas conveyed via text (Cao et al., 2020), and word embeddings can be used to track differences in the meanings of words or how they change over time (Kozlowski et al., 2019; Stoltz & Taylor, 2021).

In short, the continual growth of new computational methods available to social scientists provides a wealth of opportunity for comparative and historical research on social movements that was previously primarily the domain of protest event analysis. While we demonstrated just a few of these methods here, our findings suggest that the potential for these methods to expand our understanding of social movements, particularly the space between protest events, is vast. As we argue at the beginning of the paper, much social movement theory is rooted in a depiction of movements as contentious and challenging of institutions. Clearly, there is a portion of organizations in most movements that do in fact fit this mold. Our theories should account for the presence of local movement organizations that seek to build community, establish new institutional alternatives, and collaborate with nonmovement actors. Although many qualitative scholars and specific branches of social movement theory focused on identity building have allowed room for theorizing noncontentious movement activity (e.g., Melucci, 1980; Polletta, 2002), these types of activities are often forgotten in the building of quantitative datasets used to test hypotheses about movement activities. We believe that the methods we introduce in this chapter open up the possibility to study these types of organizations and movement activities in a more systematic way across a wide variety of movements. Although we apply our analyses to a relatively mature movement that exhibits variety in activities and organizational forms, we believe the methods shown here can be equally useful in studying movements at different stages of existence and will be especially useful as we seek

to understand how movements evolve and grow over time. This, we argue, is the future open to scholarship on social movements, conflict, and change.

NOTES

1. https://clinecenter.illinois.edu/project/human-loop-event-data-projects/SPEED. Accessed on October 14, 2021.
2. The newspapers archived by Nexis Uni and EBSCO are much more sporadic before 1990.
3. We tried to compare the prevalence of protest events across types of articles, but the section type was not consistent enough across newspapers and the two databases to do this comparison. Based on select comparisons, it is likely that news articles are more likely to report events, while lifestyle and opinion articles are less likely to report events. Future research could attempt to do this comparison more systematically.
4. https://www.infoplease.com/culture-entertainment/journalism-literature/top-100-newspapers-united-states. Accessed on October 13, 2021.
5. If the same article was published in multiple newspapers, we kept one instance of the article in each newspaper it appeared in.
6. For a full list of tactics in each category, see the online appendix: https://osf.io/preprints/socarxiv/tc2yn/.

REFERENCES

Amenta, E., Caren, N., Olasky, S. J., & Stobaugh, J. E. (2009). All the movements fit to print: Who, what, when, where, and why SMO families appeared in the New York Times in the twentieth century. *American Sociological Review*, *74*(4), 636–656.

Andrews, K., & Gaby, S. (2020). Protest episodes: Shifting actors and targets in local movements. In J. Jasper & B. G. King (Eds.), *Protestors and their targets* (pp. 124–144). Temple University Press.

Andrews, K. T., & Caren, N. (2010). Making the news: Movement organizations, media attention, and the public agenda. *American Sociological Review*, *75*(6), 841–866.

Bonilla, T., & Grimmer, J. (2013). Elevated threat levels and decreased expectations: How democracy handles terrorist threats. *Poetics*, *41*(6), 650–669.

Brulle, R. J. (2000). *Agency, democracy, and nature: The US environmental movement from a critical theory perspective*. MIT Press.

Camp, S. M. (2005). *Closer to freedom: Enslaved women and everyday resistance in the plantation south*. University of North Carolina Press.

Cao, H., Cheng, M., Cen, Z., McFarland, D. A., & Ren, X. (2020). Will this idea spread beyond academia? Understanding knowledge transfer of scientific concepts across text corpora. *arXiv preprint arXiv:2010.06657*.

Carmichael, J. T., Jenkins, J. C., & Brulle, R. J. (2012). Building environmentalism: The founding of environmental movement organizations in the United States, 1900–2000. *The Sociological Quarterly*, *53*(3), 422–453.

Chao, D. (2020). Connecting environmental justice and environmental literacy to education. Retrieved from https://tenstrands.org/environmental-literacy/connecting-environmental-justice-and-environmentalliteracy-to-education/

Corrigall-Brown, C. (2016). What gets covered? An examination of media coverage of the environmental movement in Canada. *Canadian Review of Sociology*, *53*(1), 72–93.

Della Porta, D., & Rucht, D. (1995). *Left-libertarian movements in context: A comparison of Italy and West Germany, 1965–1990*. Routledge.

Demarest, L., & Langer, A. (2022). How events enter (or not) data sets: The pitfalls and guidelines of using newspapers in the study of conflict. *Sociological Methods & Research*, *51*(2), 632–666.

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics*, *41*(6), 570–606.

Doherty, B. (1999). Manufactured vulnerability: Eco-activist tactics in Britain. *Mobilization: An International Quarterly*, *4*(1), 75–89.

Earl, J., Martin, A., Soule, S. A., & McCarthy, J. (2004). The use of newspaper data in the study of collective action. *Annual Review of Sociology*, *30*, 65–80.

EarthJustice. (2021). Landmark legal cases to build a justice-centered environmental movement. Retrieved from https://earthjustice.org/feature/landmark-cases/justice/

Evans, J. A., & Aceves, P. (2016). Machine translation: Mining text for social theory. *Annual Review of Sociology*, *42*, 21–50.

Fisher, D. R. (2019). *American resistance*. Columbia University Press.

Fisher, D. R., Andrews, K. T., Caren, N., Chenoweth, E., Heaney, M. T., Leung, T., Perkins, L. N., & Pressman, J. (2019). The science of contemporary street protest: New efforts in the United States. *Science Advances*, *5*(10), eaaw5461.

Gordon, A. D. (1996). Null models in cluster validation. In *From data to knowledge: theoretical and practical aspects of classification, data analysis, and knowledge organization* (pp. 32–44). Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-642-79999-0_3

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, *21*(3), 267–297.

Gusfield, J. R. (1970). *Protest, reform and revolt: A reader in social movements*. John Wiley & Sons.

Hanna, A. (2017a). MPEDS: Machine-learning protest event data system (v1.0). Zenodo.

Hanna, A. (2017b). MPEDS: Automating the generation of protest event data. Retrieved from https://osf.io/preprints/socarxiv/xuqmv

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949.

Lazer, D., & Radford, J. (2017). Data ex machina: Introduction to big data. *Annual Review of Sociology*, *43*, 19–39.

Leetaru, K., & Schrodt, P. A. (2013). GDELT: Global data on events, location, and tone, 1979–2012. In *ISA Annual Convention* (Vol. 2, pp. 1–49).

Lipsky, M. (1968). Protest as a political resource. *American Political Science Review*, *62*(4), 1144–1158.

McAdam, D., Sampson, R., Weffer, S., & MacIndoe, H. (2006). "There will be fighting in the streets": The distorting lens of social movement theory. *Mobilization: An International Quarterly*, *10*(1), 1–18.

McAdam, D., & Su, Y. (2002). The war at home: Antiwar protests and congressional voting, 1965 to 1973. *American Sociological Review*, *67*, 696–721.

McAdam, D., Tarrow, S., & Tilly, C. (2001). *Dynamics of contention*. Cambridge University Press.

Melucci, A. (1980). The new social movements: A theoretical approach. *Social Science Information*, *19*(2), 199–226.

Mohr, J. W., & Bogdanov, P. (2013). *Introduction—Topic models: What they are and why they matter* (pp. 545–569). Elsevier.

Molina, M., & Garip, F. (2019). Machine learning for sociology. *Annual Review of Sociology*, *45*, 27–45.

Myers, D. J., & Caniglia, B. S. (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.

Nelson, L. K., & King, B. G. (2020). The meaning of action: Linking goal orientations, tactics, and strategies in the environmental movement. *Mobilization: An International Quarterly*, *25*(3), 315–338.

O'brien, S. P. (2010). Crisis early warning and decision support: Contemporary approaches and thoughts on future research. *International Studies Review*, *12*(1), 87–104.

Oliver, P., Hanna, A., & Lim, C. (2019). Black protest events in the US 1994-2010: Issues, campaigns and trends. *Sociological Science*, *9*(12), 275–312.

Oliver, P. E., & Maney, G. M. (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, J. E., & Rahn, W. M. (2016). Rise of the Trumpenvolk: Populism in the 2016 election. *The Annals of the American Academy of Political and Social Science*, *667*(1), 189–206.

Pitt, J. (2021). Audubon's Jennifer Pitt testifies before U.S. Senate on drought and climate change. Audubon Society.

Polletta, F. (2002). *Freedom is an endless meeting: Democracy in American social movements*. University of Chicago Press.

Reger, J. (2012). *Everywhere and nowhere: The state of contemporary feminism in the United States*. Oxford University Press.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, *58*(4), 1064–1082.

Schwartz, H. A., & Ungar, L. H. (2015). Data-driven content analysis of social media: A systematic overview of automated methods. *The Annals of the American Academy of Political and Social Science*, *659*(1), 78–94.

Scott, J. C. (1985). *Weapons of the weak: Everyday forms of peasant resistance*. Yale University Press.

Snow, D. A., & Soule, S. A. (2010). *A primer in social movements*. W. W. Norton.

Staggenborg, S., & Taylor, V. (2005). Whatever happened to the women's movement? *Mobilization: An International Quarterly*, *10*(1), 37–52.

Stoltz, D. S., & Taylor, M. A. (2021). Cultural cartography with word embeddings. *Poetics*, *88*, 101567.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, *29*(1), 24–54.

Taylor, V. A. (1996). *Rock-a-by baby: Feminism, self help, and postpartum depression*. Psychology Press.

Tuğal, C. (2009). Transforming everyday life: Islamism and social movement theory. *Theory and Society*, *38*(5), 423–458.

Wang, W., Kennedy, R., Lazer, D., & Ramakrishnan, N. (2016). Growing pains for global monitoring of societal events. *Science*, *353*(6307), 1502–1503.

Whittier, N. (2010). *Feminist generations: The persistence of the radical women's movement*. Temple University Press.