This is an Accepted Manuscript of an article published in
Mobilization: An International Quarterly on 02 July 2024 (2024, Volume 29, Issue 2, pp. 245–262),
available at: https://doi.org/10.17813/1086-671X-29-2-245.

Expanding protest event analysis through videos

Donatella della Porta[a], Sophia Hunger[bc], Swen Hutter[cd], Anna Lavizzari[e]

[a] *Scuola Normale Superiore, Florence, Italy*

[b] *Universität Bremen, Bremen, Germany*

[c] *WZB Berlin Social Science Center, Berlin, Germany*

[d] *Freie Universität Berlin, Berlin, Germany*

[e] *Universidad Complutense de Madrid, Madrid, Spain*

Abstract

The increasing availability of digital video material has led to its widespread use in the social sciences, especially in research adopting processual and relational approaches. However, methodological reflection has not developed at the same pace. In this article, we propose ways to extend classical protest event analysis through the use of videos of contentious events. Our novel method, video-enhanced protest event analysis (vPEA), expands protest event research on two frontiers. First, vPEA can be used extractively, contributing to source triangulation and increasing the validity of written sources. Second, it can be used reconstructively, expanding the range of information available by capturing emotional atmosphere and choreography, as well as processes and dynamics more generally. The article provides a step-by-step guide through the vPEA process, which we illustrate using a secondary analysis of existing research and original material on Covid-related protests in Germany and Italy.

Keywords: protest event analysis, video analysis, triangulation, digital data

Introduction: Developing opportunities for video analysis

Digital technologies have rapidly increased the availability of various materials, especially video recordings. Likewise, digital-video integration has impacted the social sciences, providing students with alternative ways to disseminate research findings and innovative approaches to conducting fieldwork (Shrum et al. 2005: 1). This has generated particular interest among scholars of processual and relational approaches. However, as Nassauer and Legewie (2021) state in their insightful review, methodological reflections and a common methodological framework have not developed at the same pace (see Legewie and Nassauer 2023).

In this paper, we aim to contribute to the ongoing methodological debate about the opportunities and challenges of using video data analysis in studying social movements and contentious politics. Our aim is to offer a fresh perspective by exploring how classical protest event analysis (PEA) can be expanded using the wealth of video content now available online.[1] While PEA has traditionally focused on quantifying and characterizing large numbers of contentious events, often based on newspapers, we seek to discuss ways to combine this approach with emerging video data analysis. Ultimately, we wish to bridge the gap between the aggregative logic of classical PEA and an understanding of the situational protest dynamics by implementing video-data-analysis techniques. In doing so, we draw on previous research that has combined written and visual materials to address various aspects of protest events (see, in particular, Ketchley 2017; Grimm 2022; della Porta and Tufaro 2022; della Porta et al. 2023).

As research on social movement communication has noted, activists are increasingly using cameras (often from smartphones) to record protest events, before posting the resulting videos on

[1] For reviews on protest event analysis, see Koopmans and Rucht 2002; Earl et al. 2004; Hutter 2014; Oliver, Hanna, and Lim 2022.

social media and various platforms (e.g., Cammaerts 2015; Poell and van Dijck 2017). In addition, journalists are using more and more visuals in their coverage of protests, publishing videos of contentious events on news websites and social media platforms. At times, they also invite activists to deposit their visual material or connect with activist media in a hybrid media environment. This applies even more so to less routinized forms of protest, to large events, and to episodes of violence—which are indeed the types of protests more often covered by all types of media (Earl et al. 2004). In addition, activist-generated content is more likely to be used at events attended by younger and more technologically savvy people, who value personalized politics more, but it is also especially likely to emerge at protests in authoritarian contexts, where mass media coverage is not likely and it is all the more relevant to raise the awareness of the distant public.

Given the vast range of available sources, we suggest that video-enhanced protest event analysis (vPEA) can expand protest event research on two frontiers. First, it can facilitate a triangulation of sources, increasing the validity of written sources and enabling assessments of protest aspects hardly covered in newspaper reporting. This is what we call the *extractive function* of vPEA. Second, vPEA can also be used in a *reconstructive way* by extending the range of available information, particularly with reference to the emotional atmosphere, collective framing, event choreography, and situational protest dynamics more generally. We argue that digital ethnography (focusing on fewer events but with thick descriptions of them) or upscaling based on computational techniques (focusing on more events but covering fewer variables) can be used to fulfill both of these functions. Compared to participant observation at events, one of the traditional alternatives in protest analysis, the main advantage of vPEA is that it reduces the material costs and human resources invested while offering expanded opportunities to cover past or distant events. The price researchers pay is that they analyze materials that they do not collect first-hand and that therefore

may be less accurate or even biased by the filmers' and posters' interests and strategies. However, the analysis of existing videos can, we claim, enrich our knowledge of both single events and protest waves more generally.

In order to take advantage of these opportunities and avoid potential pitfalls, we suggest carefully assessing trade-offs in terms of advantages and challenges at each stage of the research process. On the one hand, as we will discuss, vPEA can help to mitigate problems of description bias in journalists' written reports, which are often very brief and, of course, largely mediated by considerations of newsworthiness, and by the information available to the journalists themselves. As Grimm (2022: 299) has noted in his research on Egypt, coverage on citizen reporter platforms enables researchers "to check samples for description bias and bolster the reliability of the collected data." On the other hand, however, combinations of PEA with video analysis can also introduce biases that we need to be aware of.

The paper is structured as follows: First, we provide a brief overview of protest event analysis and situate vPEA within PEA. We then provide a step-by-step, practical guide to the research process of a video-enhanced protest event analysis, from the theoretical questions that vPEA can answer, to the case and material selection, and to the data collection and analysis strategies. We illustrate our approach using secondary analysis of existing research and original material drawn from our own comparative analysis of Covid-related protests in Italy and Germany, in which we tested vPEA as a useful and versatile method for studying heterogeneous and dynamically evolving protest mobilizations.

Situating vPEA within protest event research

Social movements are about much more than just protest. Nevertheless, unconventional forms of action are a key channel through which they pressure decision-makers, become visible to the public, and generate commitment from their supporters (della Porta and Diani 2020). This is why social movement studies have paid much attention to the repertoire of protest and why PEA has become a key workhorse of research on social movements and contentious politics since the 1980s and early 1990s. Researchers have broadly relied on PEA to systematically assess the amount and features of protests across geographical areas (from the local level to the supranational level), across issue fields (from welfare to the environment), and over time (from short periods to several decades). While being aware of some biases, social movement scholars have mainly used newspaper articles as their primary textual sources.

As researchers have continuously attempted to improve PEA as a data collection technique and exploit its advantages, protest event research has proceeded through (at least) *five generations*:[2] The first generation, the "initiators"—as Rucht, Koopmans, and Neidhardt (1998:10) called them—were researchers interested in covering various indicators for a large number of countries (e.g., Taylor and Hudson 1972) or long-term processes of social and political change (e.g., Tilly, Tilly, and Tilly 1975). These authors paid limited attention to creating fine-grained categories and did not document their procedures in depth. The resulting shortcomings were addressed by the second generation of scholars, who made more extensive use of protest event data by breaking it down according to various analytic criteria and assessing the over-time development of protest and

[2] For more on the development of PEA, see also Davenport (2009:25ff), Hutter (2014, 2019), Koopmans and Rucht (2002:232ff), and Tilly (2008:19ff).

its cross-national variance[3]. Although these second-generation authors were more sophisticated in their definition of coding procedures and source selection, they did not invest much in addressing bias in their sources. The latter has become a central focus of a third, methodologically interested generation of protest event research. This generation particularly assessed the selection (what events get covered?) and description (what information gets covered?) bias of newspaper coverage (for a review, see Earl et al. 2004).

Since the rather critical debates around PEA and its bias in the late 1990s/early 2000s, the method has gained ground again, revealing its enduring popularity and flexibility. This popularity is illustrated by Figure 1, which shows the yearly number of PEA-based journal articles published since 2000 according to Google Scholar. In this period, there are two further generations of protest event research. A fourth generation applied (semi-)automated "text-as-data" approaches to speed up the selection and coding of protest events.[4] Such techniques make coding multiple sources or social media content easier. Many new datasets and initiatives were established in these years, not least due to the increasing availability of digital archives and advances in computational social sciences.[5] However, many of these datasets linked to the fourth generation of PEA research followed previous approaches in that they only counted the number of protest events. In that sense, they adopted what Tilly (2008, 26) called an epidemiological level of analysis. However, Tilly

[3] Exemplary studies include McAdam's (1982) work on civil rights protests in the United States, Tarrow's (1989) study on the Italian protest cycle in the late 1960s/early 1970s, and Kriesi et al.'s (1995) comparative perspective on new social movements in Western Europe.

[4] Early examples include the European protest and coercion data collected by Francisco and colleagues (e.g., Francisco 1996) and the study on Europeanized protests by Imig and Tarrow (2001).

[5] These include the GDELT project (https://www.gdeltproject.org/), the Armed Conflict Location & Event Data Project ACLED (https://acleddata.com/; Raleigh et al. 2023), the Observatory for Political Conflict and Democracy POLDEM (https://poldem.eui.eu/; Lorenzini et al. 2022), the Collective Action from Social Media CASM (Zhang and Pan 2019), or the Far-Right Protest in Europe FARPE dataset (Gastelli Gattinara et al. 2022).

identified two other levels: the "narrative" approach (which aims to reconstruct interactions during a contentious episode) and the "middle ground" (aiming to reconstruct interactions across multiple episodes). Those two levels have been addressed by a fifth generation of scholars, who share an ambition to understand the relational character of political contention better—be it by unpacking single contentious performances (e.g., Tilly 2008), linking different protest events to each other (e.g., Diani and Kousis 2014; Oliver et al. 2022), or assessing interactions between different actors across a set of contentious episodes (e.g., Bojar et al. 2021; Kriesi et al. 2021).

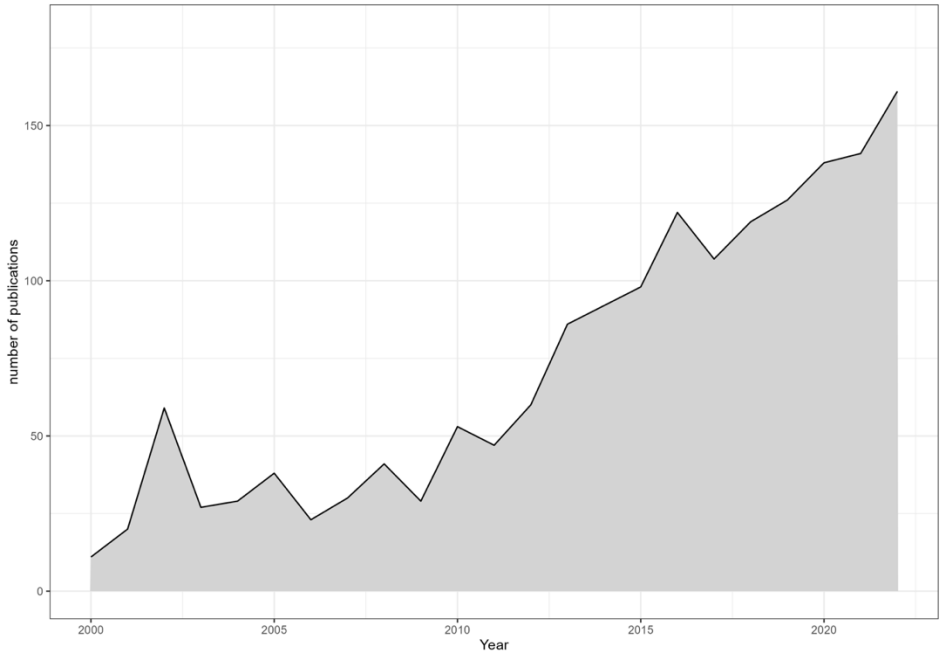

FIGURE 1 [# PEA-based articles by year]

As we argue in this paper, we regard the systematic incorporation of video material into protest event research as an opportunity to enrich classical PEA and its aggregative logic by focusing more strongly on situational dynamics. In this regard, vPEA also offers an opportunity to better connect the fourth generation of protest event research (with its emphasis on multiple

sources and big data) and the fifth generation (with its emphasis on relations and interactions among different actor types).

Theoretical questions to be addressed by a video-enhanced protest event analysis

Social movement studies often engage in the thick description of events, which are understood as key expressions of contentious politics. Focusing on processes within events, research has pointed to the importance of protests as arenas in which different players intervene and interact (della Porta 2009). Especially when looking at protests of relatively long duration, researchers have singled out different forms of action. Their sequencing is particularly relevant for understanding dynamics of escalation, but also, more generally, for grasping the processual nature of protest events (della Porta 2020).

In social movement studies, videos are regarded as important sources for addressing some key questions on the symbolic, emotional, and cognitive mechanisms in protest events, and for answering some key questions addressing their internal dynamics. In this sense, we can identify an *extractive* use of the additional information contained in videos oriented to fill in missing values and indicators not extensively covered in texts and a *reconstructive* use oriented towards analyzing dynamic processes within events (Grimm 2022).

We begin with the questions that can be addressed through a *reconstructive* use of vPEA, since we regard video analysis as particularly suited for studying situational dynamics, particularly interactions and expressions of emotions. This is also what Nassauer and Legewie emphasize in their insightful work on video data analysis (see Nassauer and Legewie 2019, 2021; Legewie and Nassauer 2023). As Nassauer and Legewie (2019:2) put it, "Analyzing situational dynamics means

to seek understanding of the rules, processes, and sequential patterns that govern social life on the micro level, both in everyday encounters and extreme situations. At the core of this perspective lies the question: How do social actions and situational dynamics impact social outcomes?" Research on violent escalation has particularly stressed how video analysis allows researchers to study *social interactions*, which involve "any action performed by someone who is motivated by, oriented to and coordinated with others, irrespective of whether these 'others' are other participants, animals, artefacts, or whatever" (Knoblauch, Tuma, and Schnettler 2013:436).

Using videos from diverse sources might help researchers to correct biases related to television and newspaper coverage of movements and movement organizations by offering an in-depth look at protest developments (Davenport 2012). For this reason, visual analysis has been connected with *visual studies* in its use of visual data, to *visual ethnography* in its focus on situational dynamics (Shrum et al. 2005), to *experimental behavioral studies* in psychology (as it is based on detailed analysis of recorded individual and social behaviors), and to *multimodal interaction analysis*, which seeks to study all relevant interactional dimensions of specific situations, including intonations, facial expressions, gestures, gazes, and body postures, with particular attention to turn-taking and communication sequences (Nassauer and Legewie 2021:138-143).

Videos are likewise good sources for the analysis of temporal dynamics, as they provide detailed information on sequences of interactions. As a *temporal* medium, i.e., a medium that records in stretches of time, video recordings allow researchers to study the structural temporal ordering of the ongoing activity. A most important advantage of videos over photography is the capacity of videos to reveal the processual dynamics of social interactions. This aim was central in the collective research project *Mapping # NoG20*, which sought to reconstruct the protest events during the counter-summit to the G20 summit in Hamburg in 2017. The use of videos allowed

researchers to carefully reconstruct all the different moments of violent escalation between July 2 and July 8, paying attention to all actors involved in situational dynamics that were embedded in both local and global contexts (Malthaner and Teune 2023). Focusing instead on emotions and memory, della Porta and Tufaro (2022) have used videos to create a thick description of some main contentious moments in Lebanon; they reflect on how the very same place in downtown Beirut was used, through different choreographies, to explain how different actors develop an emotional and cognitive mechanism and build a nostalgic narrative of the past within a movement-countermovement dynamic. Furthermore, vPEA and its reconstructive function might offer insights into the mobilizing power of protests and how protest participation can establish a shared identity between protest participants, driven by its "enjoyment value" (Hager et al. 2022), that is, music, artistic expression, and a party-like atmosphere (Danaher 2010; Eyerman 2002).

In collecting data on the Covid-19 protests in Germany and Italy, we were especially interested in reconstructing the evolving general atmosphere at protest events. In particular, we looked at the emotions being displayed, for instance, by observing whether people were chanting, applauding, or cheerfully singing or whether anger was expressed through speeches, shouting, or incitement to violence. Other parts of the analysis considered interactions , that is, how the protestors related to each other, the police, bystanders, counter-protestors, and visibly present journalists. This included a description of the nature of the interactions, for example, as tense, peaceful, or violent. The temporal sequence of these interactions was also recorded; this concerned which groups were present during the event, whether they were there from the beginning or joined later on, or whether they left at any point in time. Special attention was devoted to the changing nature of the interaction, for instance, as interactions escalated from peaceful to conflictual or violent

ones. Overall, we observed a reoccurring protest script in both countries: Initially there were peaceful protest events, often marked by provocative behavior and disregard for pandemic-containment measures, which gradually escalated into increasingly aggressive and confrontational dynamics.

Next to this reconstructive function of vPEA, we consider the *extractive use* of vPEA, as video material can offer a much richer description of the various features of protest events as symbolic actions that feature choreographies, rituals, slogans, and prefigurative acts. Information on these dimensions cannot be easily classified in the simplistic and broad categories used by PEA (such as marches, strikes, or sit-ins). vPEA hence responds to the need to analyze the changing forms of action—because, while repertoires of action tend to become standardized over time, they constantly adapt to a mutable context and to the tastes of new generations of activists. Learning processes and new affordances open up innovations at the margins. The choreography of a demonstration changes, often dramatically, impacting the symbolic dimension of protest and meaning-making among activists. To give just one example, while marches are considered a consolidated part of modern protest repertoires, performances within marches have changed dramatically, from the well-ordered, monochrome marches of the labor movements to the multicolored pride marshes, which include music and theater performances. This is all the more relevant as some protest performances have become central elements for collective identification, such as the social forum for the alterglobalist movement in the beginning of the 2000s and the "acampadas" of the anti-austerity protests at the beginning of the 2010s. As the Black Lives Matter global campaign against systemic racism and police violence indicates, some specific performances—such as the genuflection with a risen fist for the length of time of George Floyd's agony—spread crossnationally, contributing to global identification (della Porta, Lavizzari and Reiter 2022).

In addition, the extractive use of vPEA can help to mitigate the endemic missing data problem of newspaper-based PEA, as media accounts rarely contain detailed information on critical variables of interest for scholars of social movements and contentious politics. For example, screen shots of videos, or pictures more generally, can help researchers to assess the event size as measured by the number of participants, the number and type of protest organizers, and the presence of the police (e.g., Grimm 2022; Sobolev et al. 2020). Beyond filling in missing values, vPEA can help expand mainstream protest event analysis by offering further information on the organization of the space, the type of participants, the slogans, or the posters—aspects that are hardly covered systematically in news reporting. In a recent study, Bernasco et al. (2023) used computer vision based on video recordings of public places in the Netherlands to study individuals' social distancing behavior. Such techniques could also be used to compare compliance with rules at different protest events during the pandemic. Moreover, given that videos are audiovisual materials, they can enable researchers to expand their knowledge of the frames used by the speakers at events. While written texts can also be used to analyze frames, the speeches presented at protest tend to be different in style and content.

In our research on the antivax protests, we used video to expand the theoretical reach of the protest event analysis, focusing on the cognitive and the affective mechanisms of the protest events and paying particular attention to relational and processual approaches. The vPEA process means that talks can be transcribed, slogans noted, and organizational names on banners and flag systematically coded. For researchers studying regressive movements, the use of video footage captured by activists or bystanders is even more relevant, as traditional methods like participant observation or interviews may encounter challenges, especially during later stages of mobilization, when media attention wanes. Videos also provide contextual information that is richer than traditional media

reports, allowing researchers to observe participants' visible characteristics (age, gender, clothing styles) and the setting of the protest. Emotions are also often conveyed through the tone of speeches, as well as through interactions among protestors and between protesters and bystanders. While considering the protests against the Covid 19 containment measures, for example, we could in fact reconstruct in detail which slogans and heckles were chanted and which posters, flags, and symbols were displayed by participants. In the German case, for instance, our analysis of slogans, banners and speeches revealed that conspiracy theories, such as the belief that vaccination was being used to suppress citizens, were present as early as in May 2020; this was, however, not reflected by the media discourse on the emerging movement. In addition to conspiratorial thinking, in Italy we could detect a combination of nationalistic and far-right symbols, with a frequent display of Italian flags, religious references on banners calling for divine justice, and the appropriation of symbols from popular culture evoking resistance and bravery, such as masks and costumes.

Videos as data sources for vPEA

When analyzing protests, researchers can use of video to bridge newspaper-based PEA with digital ethnographic methodologies. The latter usually offer the possibility of thick descriptions—but for obvious reasons (even when multi-sited), they can cover only a very limited number of events. PEA, by contrast, usually uses media reports (or other sources of archival data) to collect (and code) static and rather "dry" factual information on a relatively large number of occurrences. Through videos, which are "captured in real time, contention appears unruly and emotional, with micro-interactions appearing to be formative in explaining situational outcomes" (Ketchley 2017:15). In fact, vPEA offers researchers options to combine breadth and depth by increasing

the qualitative information available and by reducing the costs for observing them. The thick description coming from readily available videos facilitates an in-depth analysis of a broader number of cases than is usually allowed in ethnographic research, substantially reducing the cost of collecting observations.

Furthermore, vPEA provides an avenue to study difficult-to-access protests, for instance, due to travel restrictions (as in the Covid-19 pandemic) or due to other visa, travel, or safety restrictions. This also applies in circumstances where participation might be risky for the researcher—for instance, fieldwork in potentially violent protests or in authoritarian contexts. Due to the availability of videos, scholars may also carry out in-depth analyses of past protest events. What's more, by reducing the biased effects of a filtered written coverage, video analysis can substantially expand the information collected on a large number of events, especially regarding symbols and frames. For example, the research on nostalgia as expressed in three important demonstrations in Lebanon was made possible by the extensive observation of hour-long live footage of the protests, which was shot and streamed by the main national TV channels (della Porta and Tufaro 2022).

In general, the advantages of video data are related to the potentially dense and detailed information that may be gained (Nassauer and Legewie 2019:3). To begin with, visual data represent a significant advancement in accessing information concerning real-life situational dynamics and human interactions. Second, they allow researchers to access the details of situational occurrences, circumventing potential observer bias, oversight after the event, and memory lapses. Third, visual data promote transparency, aiding researchers in openly sharing primary data on human interactions. Fourth, the analysis of existing videos saves time, money, and energy compared to participant observation at events, and also reduces biases potentially related to the visible presence

of researchers in the field. Finally, as Nassauer and Legewie (2021:142) rightly note, videos available online are particularly useful sources on *rare events,* which, given the spread of new technologies, are increasingly being photographed or recorded on video.

Although they do not allow viewers to experience smells or temperatures, videos have been called "a *social microscope* able to examine the *minutiae of ongoing social interaction* and the ways in which they contribute to the social construction of reality. In addition to the temporal order and the form of movement, video also conserves those features of situational arrangements that remain permanent throughout the course of interaction. Static artefacts, equipment, accessories, spatial organization, color, texture and the atmosphere of the immediate environment in which the interaction is taking place are depicted, at least partly conserved in two-dimensional representations and thus can later be reconstructed from the video, including its symbolic meaning" (Knoblauch et al. 2013:440).

A potential advantage of videos is that the researcher does not risk influencing an event due to their presence as an outsider. Unlike ethnographic work, digital ethnography allows researchers to replicate results by repeatedly analyzing the materials.

Of course, there is also a trade-off inherent in the use of videos, as they are more filtered than participant observation and more time-consuming than simple analysis of written coverage. The description of events will necessarily be biased by the motifs and capacity of those who film and post; some evidence might be fabricated, the content may be strategically edited,[6] and the information on the source may be insufficient to locate the action in time and space. Moreover, the very presence of video cameras might alter the sequence of the events (as might the presence of

[6] An extreme form of strategic use of videos is, for instance, the use of suicide bombers' last video-recorded messages (Straub 2016).

journalists more generally). As with other techniques, researchers using videos ought to be concerned with the reliability of the sources and their coverage. In the case of videos on protests, it is particular important to reconstruct the intentions of the video makers (Reichertz and Englert 2011); they should rely on trusted sources and/or triangulating sources (see below). In general, researchers should reflect critically on whether and how the source/s chosen might have introduced biases in terms of the reliability of the materials.

Videos should also be assessed in terms of their quality, especially their capacity to cover various aspects of the chosen event. In this regard, researchers have spoken of *optimal capture* when visual data cover the entire duration of a situation or event, its space, and all actors involved. This, of course, very rarely happens, as videos tend to cover a limited space, a limited time, and thus a limited portion of the events. Selection can be made on strategic grounds, ranging from activists' and observers' political aims to the professional concerns of the journalists but also the esthetic aspirations of the video producers (Snowdon 2020).

To improve data reliability, researchers can *triangulate* video data with other types of data, including interviews with participants (who can be asked to supplement visual data with more information and additional materials), documents, field notes from participant observation, and descriptions in the published articles on which PEA is usually based. McCluskey and Uchida (2023), for example, identify substantive and systematic gaps between police body camera footage and video-based field observation by trained coders, highlighting the potential of triangulating the two strategies of data collection. Likewise, in the research on nostalgia in social movement mobilization in Lebanon, a variety of secondary sources, including journal articles and written testimonies (della Porta and Tufaro 2022), were used to triangulate the video data. In the above-mentioned research on the anti-G20 campaign in Hamburg, the analysis of existing videos was combined with

interviews with police officers and protestors, a survey of protestors, analysis of media coverage, and document analysis (Malthaner and Teune 2023).

The availability of videos for future scrutiny increases the validity of the data as "video generated data is an ideal resource in as far as it can provide a faithful record of the process as an aspect of the naturally occurring interaction which comprises the research topic" (Lomax and Casey 1998:121). We have to consider, however, that the presence of videos on the web cannot be taken for granted. Hence, in order to insure further use and checks, videos must be downloaded and safely stored. Scholars have no control over videos that might be taken down—by the producer itself or through deplatforming. Additionally, the same images might be decoded in different ways by different observers. As with participant observation, this requires researchers to be self-reflexive and, as with PEA, it means that the test of intercoder reliability should be conducted when multiple researchers participate in coding.

Another concern relates to time investment. Collecting detailed accounts of protest events is no trivial exercise; rather, it can be a time- and labor-intensive process. For those seeking to conduct large-N studies, data storage is no trivial exercise. Moreover, viewing an hour of video while recording the characteristics of interest can result in up to two to three hours of work. Generally, collecting dense data requires the researcher to pause often and rewind the video, especially in situations of escalation or in shots where many banners are visible. Furthermore, analyzing speeches given at rallies is labor intensive, since the researcher has to transcribe large shares of the statements made and categorize the claims and frames immediately. By contrast, some videos that broadcast demonstrations from an individual's point of view may not be all that labor intensive as they contain more extended parts in which little happens, except for people marching. Also, the

amount of time invested varies a lot according to the number of events covered and, of course, the length of the video coverage.

There are various ethical concerns (see Legewie and Nassauer 2018). While videos are often easily accessible, video analysis presents specific ethical challenges. The first question is the *legality* of access. In this regard, it is important to consider how the platforms on which videos are published are managed, including any restrictions on the use of the information provided. In addition, privacy concerns should be addressed, especially regarding the expectations and motivations of those who upload video and the public (or private) nature of the events and spaces covered. Concerning *anonymity*, "a study is ethically less problematic if personal characteristics of people as well as of places are indiscernible, if additional information on people and places is difficult to access if captured behavior is mundane and (if criminal behavior is analyzed) if data would likely be used for prosecution" (Legewie and Nassauer 2018: para. 36). In times of disinformation and deep fakes, researchers also need to be concerned with how reliable the information provided is.

When analyzing videos on protests, besides observing the general principle of doing no harm, researchers can reduce ethical concerns related to legal issues and confidentiality if they work with publicly available documents. While protests are public events, expectations about an individual's right to be forgotten should be considered when archiving videos. As for concerns with anonymity, these can also be reduced if the analysis covers actions carried out in public spaces. In our own research on the antivax protests, we considered only published materials available online.

Another issue to consider with vPEA is the researcher or research assistant's situation and well-being while coding. Depending on the nature and development of the protest event, research-

ers may be exposed to disturbing material in the videos, notably episodes of violence against people, clashes with the police, or even verbal aggression and hate speech, etc. In dealing with such issues, it is helpful to consider each person's sensitivity and positionality towards specific material and to plan several sessions to watch the videos (including breaks). It may even be necessary to alternate them with less challenging content. Feelings of alienation can be diffused and content can be put in the broader context of the research if researchers discuss the material perceived as disturbing with colleagues and mentors as well as taking notes, not only about the needed information but also about their own emotions (see also Ellis 2018; Grimm et al. 2020).[7]

In sum, extracting in-depth information from videos facilitates extensive and theoretically enriched research on protest events and movements. However, caution is required when using such materials: they must be assessed for potential biases, supplemented with additional information, and triangulated with other sources. As Jannis Grimm (2022:297) aptly noted,

> "In addition to providing orientation, the triangulation of methods also facilitates the verification of results, increasing the overall analytic payoff. Each level of analysis informs the other and thus contributes to an integrated understanding of the process under investigation and the reduction of potential sources of error. Time series may serve as a point of departure by drawing attention to moments of visibility and latency and to turning points in mobilization processes. In turn, close qualitative analysis of a limited set of events may provide more nuanced insights and explain transitions and nonlinearities that have been discovered through quantitative analysis."

[7] The journalistic profession has broadly discussed ways to deal with disturbing images on video and offered practical mitigating solutions (see, e.g., https://dartcenter.org/resources/online-abuse-self-defence-guide).

A case in point is Anna Schwenck's (2023) analysis of pandemic-skeptical symbolism at protests in Germany, in which she combines video clips and screenshots extracted from public pandemic-skeptical social media platforms popular in Berlin with a wide variety of sources, including participant observation at protest events and in-depth interviews. Through these multiple research sources, she reconstructs the ways in which pandemic-skeptical protesters engage in a sort of world building, that is

> "the creation of a fictional setting that is akin to a secondary world, one that operates differently from the primary world 'we' live in. Rank-and-file protesters can easily become co-creators of this secondary, often conspiracist world: they may supply novel metaphors, allegories, tropes, or stories, and they can act as multipliers of their own contributions to a growing 'alternative information' base by starting online media platforms. The protests thus thrive not only on conspiracist claims, such as that the pharmaceutical-industrial complex is behind the proclamation of the covid-19 pandemic because such a pandemic serves their financial interests, but on the creation of a pandemic-skeptical symbolic universe" (ibid. 39–40).

Case selection and data sources

A first observation is that videos on protest abound, and their availability seems destined to increase given the growing access to smartphones and use of social media in which visual content is easily uploaded. This abundance offers obvious advantages, but it presents challenges that researchers should be aware of.

First of all, complete coverage is not available on any protest event—far from it—and the posted videos often cover only small portions of the event and limited spans of time. As with PEA,

we might assume that the amount of available material tends to increase with the size, the disruptiveness, and the innovative character of each event, as well as with the contingent relevance of a given social movement or social movement organization. In addition, the type of participants—their attention to ICT and their propensity to engage in personalized forms of political participation—might affect the materials filmed and posted. Some actions (especially violent interactions) can be omitted as their publication would be risky for the activists involved (even if, as we noted, narcissism often wins over caution, meaning that even the police use videos posted by activists to reconstruct illegal actions and punish perpetrators).

When selecting cases, researchers must obviously consider the availability and quality of videos, in particular regarding the feasibility of data collection and the openness of sources. In general, the materials available vary a lot by:

a) Length, from a few minutes to full coverage;
b) Origins, from activists to external observers;
c) Content, including variable mixes of text and images;
d) Quality, regarding the professionality and technical equipment with which the recording was created.

Given the amount of available material, the quality of the material, and, of course, the main research questions, researchers can choose different sampling strategies. For ethnographic studies of social interactions, an intensive strategy might be more appropriate, while a more extensive strategy may well complement a protest event analysis based on newspapers, by adding information on a specific event that is not available in that PEA's source. In the first instance, the researcher will usually try to reconstruct the entire development of one or more critical events as

much as possible, also by combining different perspectives. This is important when aiming to explain escalation during large and complex events—such as the G20 summit in Hamburg and its counter-summit (Malthaner and Teune 2023).

In the second instance, a different strategy is used; researchers also collect short videos on a larger sample of protest events. In this case, the aim is to develop a thicker description of those events by adding information on symbolic aspects and contextual characteristics that the newspaper reports tend to overlook. Even short videos might contain data about the participants, the event choreography, the slogans, the posters, and the frames encompassed in speeches (which can be transcribed and analyzed) as well as the characteristics and general atmosphere of the space where the protest takes place. In this case, the selection can be driven by existing PEA databases or be constructed autonomously using other types of sources. Additionally, the source of the video can also become an important variable in its own right; it can indicate who actually shares information on an event or a specific feature of an event.

In general, it has been suggested for case selection that, when there are many videos on one or similar events, the criteria used to choose the ones to analyze should include the reliability of the sources and the optimal coverage of the events. Other criteria are the importance of some contexts or actors involved. When videos are short and come from less reliable sources, the use of multiple videos of the same event is advisable. One particularly relevant factor for the quality of data is the *capture*, that is "the degree to which the data material provides information on all aspects of a situation that are relevant to a given research question. In a nutshell, optimal capture means that visual data must enable researchers to establish a seamless sequence of relevant factors, and provide compelling empirical evidence for systematic links between those factors and the outcome under investigation (if such links exist)" (Nassauer and Legewie 2019:8).

In our own research on the antivax protests, we used the Internet as a source, drawing on mainstream media and social media and both activists' and third-party sources. We tested two sampling strategies. One aimed to reconstruct large and important events that represented turning points (for instance, events that are characteristic of a certain period of contestation as identified by classical PEA) by collecting information from several sources and viewpoints (for a related strategy on the combination of PEA and discourse analysis, see Grimm 2022; on the combination of videos with participant observation, see Schwenk 2023). The other strategy was based on pooling short videos on a broader number of protest events that could be regarded as representative of different periods or different protest actors. In general, the choice between these two options also relates to the two potential ambitions of vPEA—that is, their extractive versus reconstructive aim. It may be more advantageous to cover a larger number of events when pursuing the former strategy; for the latter, in-depth analysis of a few relevant events is needed. We also noticed on our vPEA that the video content may be subject to editing, meaning that frames in which escalation happens may have been cut out. Whenever these cuts were observable to the coder, we noted this in our observation template.

What to analyze

Given the variety of video material available, the potential uses of videos are also *diverse*: There are videos produced by researchers, either to document a situation or to disseminate research results, but also videos produced by others. In the latter case, researchers may ask individuals or groups to film their activities or use videos made for different purposes, often posted online. A growing number of studies have used videos available online. In such cases, analyses are influenced by the conditions under which video materials are produced, that is, "the practices by which

videos are being recorded, made accessible and handled by the researchers" (Knoblauch et al. 2013:436). In particular, attention has been paid to so-called "native videos," which proliferated as camcorders and smartphones spread. These are videos which "have been recorded – and produced in a more or less professional way – by actors other than researchers, for example private tapes made available to the researchers, video clips uploaded on YouTube or 'video diaries' triggered by the researchers" (Knoblauch et al. 2013:437).

As mentioned, videos can be used to reconstruct social interactions within events and/or for a thicker description of the events themselves. In the first instance, what the coder observes is the facial expression, the body language, the density, and the intensity of the interaction within sequential analysis. In the second case, the researcher aims to complement PEA with vPEA as much as possible, adding information on the forms of protest, the organization involved, the targets, and the claims by coding the various symbols, participant characteristics, protest spaces, emotional displays, speeches, and slogans.

In both cases, the coding may be standardized to different degrees. When analyzing social interactions, attention is especially paid to the sequences of the interactions. For this, an observation template can be devised to capture the characteristics of participants, symbols, emotions, and context. This information can also be collated with other data from published articles in newspapers. In our research on anti-pandemic-containment protests, we used video analysis on materials posted online about collective events by either the actors themselves or other observers. Moreover, we triangulated the video material with protest event analysis, which helped us contextualize the events we studied with vPEA while also serving as a guide to our sampling and case selection. Of course, more computational approaches offer the highest standardization because they capture reoccurring features such as faces, candles, or placards (e.g., Scholz et al. 2023).

In analyses of situational dynamics, facial expressions, body postures, emotions, and contexts are the most important aspects when seeking to explain interactions or more vividly describe an event. Facial expressions and body posture are especially noteworthy as nonverbal information sources that include expressions of emotions, body positions, and gestures. Interactions are expressed in verbal and nonverbal forms of communication; this includes movements, the use of objects, and ways of talking. As for the context, the characteristics of both the space in which the event unfolds and the individuals participating in them are relevant (Nassauer and Legewie 2019:5).

Regarding the coding of *facial expressions*, Collins (2016) has suggested looking in particular at the mouth and jaw, the region around the eyes, and the forehead to single out "emotions (i.e., fear, anger, happiness, sadness, surprise, disgust, and contempt), positive or negative affect, degrees of interest and engagement, and attentional foci in a situation." In addition, research has regarded aspects of body posture, such as the position of the head, shoulders, torso, arms, and legs, as indicators of emotional states. So, "relaxed body postures are characterized by shoulders down, relaxed muscles, and soft movements. They can indicate that actors feel at ease, safe, and/or self-assured, while tense body postures can indicate that actors feel threatened, uncomfortable, and/or stressed" (Nassauer and Legewie 2021:146). Focusing on vocal pitch in the United States Congress, Dietrich et al. (2023) recently presented an automated approach to data collection, highlighting how emotional intensity varies with commitment to a cause and might ultimately also influence co-legislators' behavior. Adapting this to the study of protest events could yield interesting insights into speakers' emotional intensity and the related responses by protest participants.

When looking at *interactions*, researchers have considered density, intensity, timing, and frequency. *Density* includes gestures such as nonverbal codes for communicating attitudes but also

verbal communication, from content to forms, including intonation and tone of voice. Interactions also vary as far as their *intensity* is concerned. The *timing* refers to the location of a specific action within a process of interaction. The frequency measures the number of occurrences of a certain type of interaction (Nassauer and Legewie 2021:147-48).

In addition to being relevant for assessing the situational dynamics of an event (especially escalation and de-escalation), we argue that vPEA enables scholars to assess the group dynamics of crowds at protest events, that is, whether the general atmosphere is calm and moving or heated and aggressive. Furthermore, we can grasp *whether* and *how* protestors interact with each other and others, such as bystanders and the police, that is, whether they are communicative and friendly with each other. We argue that these insights might provide useful information on escalation dynamics that might occur at a latter point of a given event.

When thinking about the context and the stable characteristics of an event, researchers may be able to extract details that remain stable over time, even from short videos. These *permanent* visual elements include "material objects and artefacts, bodily configurations and built spaces, furniture, walls, buildings, instruments, clothes, spectacles or hairstyle. These elements constitute the aforementioned 'density' of audiovisual records" (Knoblauch et al. 2013:444). The *context* of a situation includes physical and social aspects, as the physical and constructed characteristics of space gives meaning to words and deeds. Physical properties of the environment refer to

> "the expanse of space (e.g., ample or confined), lighting (e.g., bright or dark), weather conditions, and access (e.g., restricted or open). …, inanimate objects, such as tools and devices, TV screens, or audio speakers, often play a role in how a situation develops because they mediate actions and interactions between people … assignment of space for protest routes or no-protest zones is crucial to the unfolding of situational dynamics. … objects such as

police communication devices and special audio speakers developed for crowd management play a role in how demonstration marches unfold." (Nassauer and Legewie 2021:148-49).

A social dimension refers to the type of actors present in a location and their social roles.

Compared to quantitative content analysis, such as classical PEA, in vPEA, an observation template may offer researchers more leeway to grasp nuances without having to rely on pre-defined, narrow categories. At the same time, template-guided data collection may result in more structured information than can be gained through the materials usually coming from ethnographic methods such as participant observation. This may allow for easier comparison across multiple cases. In our own analysis of the Covid-19 protests, we used an observation template to collect information on some of these main dimensions. The template included the characteristics of the event (protest forms, choreography), of the participants and other actors (e.g., the police, bystanders, journalists, and counter-protestors), and of the interactions of these groups; it also covered the fluctuation in participant numbers, the verbal content of speeches and slogans, and the presence of posters, flags, and other symbols (see coding template in the appendix). Thus, we were able to extract protest characteristics, such as the relatively old age profile of protesters in Germany, and could reconstruct the dynamics of the events, which, both in Italy and in Germany, often started out as cheerful and rather easy-going events but quite often turned rather grim and radical over the course of the day.

Analyzing the data

Compared to physical forms of participant observation, video recordings have the advantage of allowing for dense description. *Density* refers to the "minute aspects (the perceptual features of

certain things, the exact course of certain events), which might have passed unnoticed by the actors or the observer in the situation, become accessible in the recorded data. …. Similar to photography and despite their constructed character, the density of video recording is enhanced by the possibility of freezing images or enlarging audiovisual frames (Knoblauch et al. 2013:439). Furthermore, researchers can delve into finer aspects by rewinding videotapes or meticulously studying photographs. This method permits them to focus on even the briefest information, like micro-expressions of emotions. It enables researchers to extract interaction sequences frame by frame, allowing them to understand extensive and intricate situations (Nassauer and Legewie 2021:141). Moreover, by replaying videos multiple times, researchers can capture the temporal evolution of interactions. Techniques such as slow or fast motion manipulation are employed to alter the recorded temporality (Knoblauch et al. 2013:440).

The aim of this situational analysis is often to reconstruct different *moments* by breaking down events into lower-level actions consisting of extremely specific elements such as shouts or shrugs. These lower-level actions can then be grouped together in higher-level actions (for instance, a conversation or a fight) via multimodal transcription, which enables researchers to link lower-level actions in chronological order. Unlike PEA, which tends to be static, in vPEA, the temporal sequences are most important thing; trey reconstructed in detail through the use of a tempo-spatial matrix of events or diagrams of movements. As in *process tracing*, in situational analysis, processes and events are reconstructed in order to link situational dynamics with outcomes; however, here, the focus is on the microlevel of situations.

Similar to other data sources and in line with the above-mentioned ways to code the information from videos, video data use varies in terms of standardization and the degree of emphasis placed on interpretation. Therefore, Knoblauch et al. (2013:436) distinguish between *standardized*

and interpretive video analysis. The former, at times using automated classification, entails systematically coding parts of the videos according to predefined categories based on theoretical assumptions. These can relate to the various features highlighted above: for example, the different emotions evoked by the various speakers. Interpretative video analysis, by contrast, assumes that "the actions recorded are guided by meanings that must be understood by the actors themselves. It is only on the basis of the meanings of actions to the actors involved, that is, 'first order constructs', that researchers pursue their questions and create their 'second order constructs'" (Knoblauch et al 2013:436).

As for other methods, data can be analyzed inductively or deductively. In inductive analysis, we have a first phase of intensive, detailed labeling of information in very short units of time and a second phase in which labels are grouped into broader categories. This iterative process can help researchers to develop a coding scheme from the data. When researchers aim to reconstruct situational dynamics, a very detailed procedure is advised. To look in particular at dynamics of escalation, Nassauer and Legewie (2019:10-11) suggested the following procedure:

> "In the first phase, the analysis we suggest researchers go through a selected set of data pieces second by second. They can label what they see in the first frame of the video (e.g., actors, objects, characteristics of the physical space) and then go on to label every change they detect in the data, e.g., regarding actors' facial expressions and posture, movements they make, or speech. These labels should be as descriptive as possible, including as little interpretation as possible. In the second phase, researchers can look at the labels they used to see whether there are overarching categories and dimensions that two or more labels fit into. For instance, 'hitting', 'punching', and other labels could be subsumed in the category 'escalatory behavior'. Once concepts have been distilled from the data in this manner, a

coding scheme derives naturally from the concepts' labels, or indicator-level dimensions, which can then be used as codes to assign to the data. As with other data types, this first phase of developing a coding scheme can be complex and rests on the researchers' knowledge and interpretation of the case and context."

In deductive approaches, which are formulated based on existing theories and previous research, concepts typically predate the data collection phase. For instance, when videos are used to expand protest event analysis, researchers can build upon existing (more or less formalized) observation templates. Similarly, machine learning approaches to protest analysis, which are commonly employed in the field of computational social science, depend on existing precoded and thus classified training data. In our own research on the antivax protests, we built our observation template around the main dimensions used for PEA on the *when*, *where*, *who*, and *how*. In addition, when present, we analyzed longer speeches for diagnostic, prognostic, and motivational frames (see appendix) (Entman 1993; Benford and Snow 2000). These different frames did play a major role in our analysis of banners and slogans shown at the events and of the speeches that were given at rallies and demonstrations. Based on previous studies on the antivax movements in Germany and Italy (della Porta and Lavizzari 2022;Hunger et al. 2023), we had already adjusted our observation template to include different problem descriptions. Some were more general, for instance, the lack of democracy in the country in question, while some were specifically tied to conspiracy theories, such as the belief in chemtrails. We also included statements on their own group as protesters (i.e., identity frames) and statements on outgroups (i.e., oppositional frames, for instance "the majority," "the government," or "the police"). Lastly, we included motivational frames in our analysis, which were often tied to calls for action, be it for further protest mobilization or more radical calls for active resistance.

In general, this form of analysis considers turning points, defined as significant alterations in the relational dynamics and emotional characteristics of the event. It examines the immediate triggers and effects of these shifts, including instances of escalation and de-escalation and actor aggregation or fragmentation. To further refine the analysis, researchers should consider the intensity in terms of the number of participants, the duration of specific interactions, and the frequencies of certain behaviors, including aggression and violence. Additionally, by focusing on specific groups, such as more violent groups of protesters or the police, researchers can gain insights into the event's evolution over time. The physical and symbolic characteristics of the space are relevant for understanding the meaning and development of an event. This not only includes the physical features of a space (or the multiple spaces) in which a protest takes place, but also its symbolic meaning (see della Porta and Tufaro 2022) and the presence of inanimate objects. As Pink (2013) observed, the visual can be combined with *other senses* in a multisensory approach.

From the technical point of view, the coding of visuals can be helped by the use of software packages; however, they are limited in terms of the dimensions they can capture. Geospatial data can also be used to aid the analysis of specific events (Grimm 2022: 299). Other scholars have suggested using automated analysis to assess how violent a protest is (Torres forthcoming) or to classify different forms of action based on visible objects (Scholz et al. 2023) via machine learning techniques. Computational methods can also be used to estimate crowd size (for an overview, see Aziz et al 2018).

In our own research on the Covid-19 protests, we collected information that allowed us to better understand sequences (including escalation) and content (especially conspiratorial thinking), and we also added observations on the mood/emotional changes, from cheerful to rather grim (even without apparent escalation in the forms of action). As emphasized before, one commonality found

in many demonstrations across both countries was the gradual change from a relaxed cheerful atmosphere to a tense and suspenseful one. Often, participants would dance and play musical instruments in the beginning, singing songs such as Queen's "I Want to Break Free." As time progressed, emotions tended to become more intense, with aggressive whistling and chanting; this was also reflected in the music, for instance "Another Brick in the Wall." This change of atmosphere was often tied to the police prematurely terminating the event or a spontaneous (and not permitted) march starting after the official rally was over. Thus, interactions with the police varied between grateful ones with police officers, presented as our "Freund und Helfer" (friend and helper), and total hostility, often at the same event. One from the anti-pandemic-containment demonstration in Berlin on August 1, 2021, illustrates this especially well: The demonstrators shouted at the police ("Let us through!"); the police prepared pepper spray. At one point, a demonstrator shouted at the police: "Are you people or machines?" Another shouted: "Then explain to your children that you took part," and "Take off your helmet and join us! We are the good guys!" Another demonstrator shouted at a policeman: "He's vaccinated his children twice!" In contrast, there was a grave atmosphere at other, mostly smaller events, where people lit up candles and speakers cried or were close to tears, asking for silence during their speeches rather than for applause. Other moments of tensions were also observable in interactions with bystanders, who were sometimes insulted or physically assaulted by protestors when they disagreed with the content of the protests. During a protest in Milan, we observed bystanders showing their Covid certs on their cellphones to protest participants or throwing water at protesters and getting insulted or assaulted.

Concluding remarks

As the amount of available digital video material has increased, so too has its use in social science research, particularly by scholars adopting processual and relational approaches. Despite this trend, there has been a lag in methodological reflection, which has hindered it from reaching its full potential. Against this backdrop, we have proposed a fresh perspective, called video-enhanced protest event analysis (vPEA). We have described ways to systematically expand on classical protest event analysis—a key workhorse of social movement studies—so that it can benefit from the contemporary information environment. More specifically, by incorporating videos of contentious events, vPEA can be used in two ways. First, it can be used in an *extractive* manner: It can contribute to source triangulation, enhance the validity of written sources, and provide systematic information on aspects hardly covered by textual sources. Second, it can be used in a more *reconstructive* manner: It can expand the range of information available by focusing on features such as emotional atmosphere, choreography, and dynamics of contestation more generally.

The paper has thus provided a step-by-step guide to the vPEA research process, with the aim of contributing to methodological reflection by providing practical information and illustrating the approach with a secondary analysis of existing research and of original materials from Covid-related protests in Italy and Germany. Ultimately, we regard the systematic incorporation of video analysis into protest event research as an opportunity to enrich both classical event counts and the more relational approaches to contentious politics. Our proposal might hence offer an opportunity to better connect ongoing efforts to construct large-scale protest event datasets via computational methods with ongoing discussions emphasizing the need for more relational protest event datasets. Of course, we have only taken a first—but hopefully crucial—step in the further development of

protest event analysis, which could enable the building of bridges between different research communities and methodological approaches.

References

Aziz, M. W., Naeem, F., Alizai, M. H., & Khan, K. B. (2018). Automated Solutions for Crowd Size Estimation. *Social Science Computer Review,* 36(5), 610-631.

Balsinger, Philip and Alexander Lambelet. 2015. "Participant Observation." Pp. 144-72 in *Methodological Practices in Social Movement Research,* edited by D. della Porta. Oxford: Oxford University Press.

Benford, Robert D., and David A. Snow. 2000. "Framing Processes and Social Movements: An Overview and Assessment." *Annual Review of Sociology* 26:611-39.

Bernasco, Wim, Evelien M. Hoeben, Dennis Koelma, Lasse Suonperä Liebst, Josephine Thomas, Joska Appelman, Cees G. M. Snoek and Marie Rosenkrantz Lindegaard. 2023. "Promise Into Practice: Application of Computer Vision in Empirical Research on Social Distancing." *Sociological Methods & Research* 52(3): 1239-1287.

Bojar, Abel, Theresa Gessler, Swen Hutter and Hanspeter Kriesi. 2021. Contentious Episodes in the Age of Austerity: Studying the Dynamics of Government-Challenger Interactions. Cambridge: Cambridge University Press.

Cammaerts, Bart. 2015. "Social Media and Activism." Pp. 1027-34 in *The International Encyclopedia of Digital Communication and Society*, edited by R. Mansell and P. Hwa Ang. Oxford: Wiley-Blackwell.

Collins, Randall. 2016. "On Visual Methods and the Growth of Micro-Interactional Sociology: Interview to Professor Randall Collins (with Uliano Conti)." Pp. 81-94 in *Lo Spazio Del Visuale. Manuale Sull'utilizzo Dell'immagine Nella Ricerca Sociale*, edited by U. Conti. Rome: Armando Editore.

Castelli Gattinara, Pietro, Caterina Froio and Andrea L. P. Pirro. 2022. "Far-right protest mobilisation in Europe: Grievances, opportunities and resources." *European Journal of Political Research* 61(4): 1019-1041.

Danaher, William F. 2010. “Music and Social Movements.” *Sociology Compass* 4 (9): 811–23.

Davenport, Christian. 2009. “Regimes, Repertoires and State Repression.” *Swiss Political Science Review* 15(2):377-85.

Davenport, Christian. 2012. *Media Bias, Perspective and State Repression*. Cambridge: Cambridge University Press.

della Porta, Donatella, and Mario Diani. 2020. *Social Movements: An Introduction.* 3rd ed. Wiley-Blackwell.

della Porta, Donatella, and Rosanna Tufaro. 2022. “Mobilizing the Past in Revolutionary Times: Memory, Counter-Memory and Nostalgia During the Lebanese Uprisings.” *Sociological Forum* 37(S1):1387-1413.

della Porta, Donatella, Anna Lavizzari, and Herbert Reiter. 2022. “The Spreading of the Black Lives Matter Movement Campaign: The Italian Case in Cross-National Perspective.” *Sociological Forum* 37(3):700-21. https://doi.org/10.1111/socf.12818

della Porta, Donatella, and Rosanna Tufaro. 2022. “Mobilizing the Past in Revolutionary Times: Memory, Counter-Memory and Nostalgia During the Lebanese Uprisings.” *Sociological Forum* 37(S1):1387-1413.

della Porta, Donatella, and Mario Diani. 2020. *Social Movements: An Introduction.* 3rd ed. Wiley-Blackwell.

Diani, Mario, and Maria Kousis. 2014. “The Duality of Claims and Events: The Greek Campaign against Troika’s Memoranda and Austerity, 2010-2012.” *Mobilization: An International Quarterly* 19(4):387-404.

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.

Ellis, Hannah. 2018. “How to Prevent, Identify and Address Vicarious Trauma — While Conducting Open Source Investigations in the Middle East.” *Bellingcat*. (https://www.bellingcat.com/resources/how-tos/2018/10/18/prevent-identify-address-vicarious-trauma-conducting-open-source-investigations-middle-east/).

Entman, Robert M.1993. “Framing: Toward Clarification of a Fractured Paradigm.” *Journal of Communication* 43(4):51–58.

Eyerman, Ron, and Andrew Jamison. 1998. *Music and Social Movements: Mobilizing Traditions in the Twentieth Century*. Cambridge: Cambridge University Press.

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

Grimm, Jannis. 2022. *Contested Legitimacies Repression and Revolt in Post-Revolutionary Egypt.* Amsterdam: Amsterdam University Press.

Grimm, Jannis, Kevin Koehler, Ellen M. Lust, Ilyas Saliba, and Isabell Schierenbeck. 2020. *Safer Field Research in the Social Sciences: A Guide to Human and Digital Security in Hostile Environments*. London: Sage.

Hager Anselm, Hensel Lukas, Hermle Johannes, Roth Christopher. Group Size and Protest Mobilization across Movements and Countermovements. American Political Science Review. 2022;116(3):1051-1066. doi:10.1017/S0003055421001131

Hunger, Sophia, Hutter, Swen, and Eylem Kanol. 2023. "The Mobilisation Potential of Anti-containment Protests in Germany." *West European Politics* 46(4):812–840.

Hutter, Swen. 2014. *Protesting Culture and Economics in Western Europe*. Minneapolis: the University of Minnesota Press.

Hutter, Swen. 2019. “Exploring the Full Conceptual Potential of Protest Event Analysis.” *Sociological Methodology* 49(1):58-63.

Imig, Douglas R., and Sidney G. Tarrow. 2001. *Contentious Europeans: Protest and Politics in an Emerging Polity*. Rowman & Littlefield Publishers.

Ketchley, Neil. 2017. *Egypt in a Time of Revolution*. Cambridge: Cambridge University Press.

Knoblauch, Hubert, René Tuma, and Bernt Schnettler. 2013. “Video Analysis and Videography.” Pp. 435-49 in *The SAGE Handbook of Qualitative Data Analysis*, edited by U. Flick. SAGE Publications Ltd.

Koopmans, Ruud, and Dieter Rucht. 2002. “Protest Event Analysis.” Pp. 231-59 in *Methods of Social Movement Research*, edited by B. Klandermans and S. Staggenborg. University of Minnesota Press.

Kriesi, Hanspeter, Ruud Koopmans, Jan W. Duyvendak, and Mario Giugni. 1995. *New Social Movements in Western Europe*. Minnesota: University of Minnesota Press.

Kriesi, Hanspeter, Jasmine Lorenzini, Bruno Wüest, and Silja Hausermann. 2020. *Contention in Times of Crisis: Recession and Political Protest in Thirty European Countries*. Cambridge: Cambridge University Press.

Legewie, Nicolas, and Anne Nassauer. 2018. “YouTube, Google, Facebook: 21st Century Online Video Research and Research Ethics.” *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* 19(3).

Legewie, Nicolas, and Anne Nassauer. 2023. "Current and Future Debates in Video Data Analysis." *Sociological Methods & Research* 52(3), 1107-1119.

Lomax, Helen and Neil Casey. 1998. “Recording Social Life: Reflexivity and Video Methodology.” *Sociological Research Online*. Retrieved May 27, 2016 (http://www.socresonline.org.uk/3/lomax/lomax_doc.html).

Lorenzini, Jasmine, Kriesi, Hanspeter., Makarov, Peter, and Bruno Wüest. 2022. "Protest Event Analysis: Developing a Semiautomated NLP Approach." *American Behavioral Scientist* 66(5), 555-577.

Malthaner, Stefan, and Simon Teune, eds. 2023. *Eskalation. G20 in Hamburg, Protest und Gewalt*. Hamburg: Hamburger Edition.

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

McCluskey, John D. and Craig D. Uchida. 2023. "Video Data Analysis and Police Body-Worn Camera Footage." *Sociological Methods & Research* 52(3): 1120-1154.

Nassauer, Anne, and Nicolas Legewie. 2019. “Analyzing 21st Century Video Data on Situational Dynamics—Issues and Challenges in Video Data Analysis.” *Social Sciences* 8(3):100.

Nassauer, Anne, and Nicolas Legewie. 2021. “Video Data Analysis: A Methodological Frame for a Novel Research Trend.” *Sociological Methods & Research* 50(1):135-174.

Oliver, Pamela, Alex Hanna, and Chaeyoon Lim. 2022. “Constructing Relational and Verifiable Protest Event Data: Four Challenges and Some Solutions.” *Mobilization: An International Quarterly*.

Pink, Sarah. 2013. “Advances in Visual Methodology: An Introduction.” Pp. 3-16 in *Advances in Visual Methodology*, edited by S. Pink. London: Sage

Poell, Thomas, and José van Dijck. 2017. “Social Media and New Protest Movements.” Pp. 546-61 in *The SAGE Handbook of Social Media*, edited by J. Burgess, A. Marwick, and T. Poell. London: Sage.

Raleigh, Clionadh, Roudabeh Kishi and Andrew Linke. 2023. "Political instability patterns are obscured by conflict dataset scope conditions, sources, and coding choices." *Humanities and Social Sciences* Communications 10(1): 74.

Reichertz, Jo, and Carina Jasmin Englert, eds. 2011. *Einführung in die qualitative Videoanalyse. Eine hermeneutisch-wissenssoziologische Fallanalys*e. Wiesbaden: VS Verlag.

Rucht, Dieter, Ruud Koopmans, and Friedhelm Neidhardt. 1998. “Introduction: Protest as a Subject of Empirical Research.” Pp. 7-30 in *Acts of Dissent. New Developments in the Study of Protest*, edited by D. Rucht, R. Koopmans, and F. Neidhardt. Berlin: Edition Sigma.

Scholz, Stefan, Nils B. Weidmann, Zachary C. Steinert-Threlkeld, Eda Keremoglu1, and Bastian Goldlüke. 2023. Improving Computer Vision Interpretability: Transparent Two-level Classification for Complex Scenes. Paper presented at the Conference " Analyzing protest in the digital age Challenges and opportunities in combining text and video sources", Berlin December 2023.

Schweingruber, David, and C. McPhail. 1999. “A Method for Systematically Observing and Recording Collective Action.” *Sociological Methods & Research* 27(4):451–98.

Schwenck, Anna. 2023.” Performances of Closeness and the Staging of Resistance with Mainstream Music Analyzing the Symbolism of Pandemic Skeptical Protests.” *German Politics and Society* 41 (2): 35–60

Shrum, Wesley, Ricardo Duque, and Timothy Brown. 2005. Digital Video as Research Practice: Methodology for the Millennium. *Journal of Research Practice* 1(1), Article M4. Retrieved [January 30, 2023], from http://jrp.icaap.org/index.php/jrp/article/view/6/11

Snowdon, Peter. 2020. *The People Are Not Images. Vernacular Video after the Arab Spring*. London: Verso.

Sobolev, Anton, M. Keith Chen, Jungseock Joo and Zachary C. Steinert-Threlkeld. 2020. "News and Geolocated Social Media Accurately Measure Protest Size Variation." *American Political Science Review* 114(4): 1343-1351.

Straub, Verena. 2016. “Suicide bomber's video testimonies.” *Nomoi*, 30 December, http://nomoi.hypotheses.org/344

Tarrow, Sidney. 1989. *Democracy and Disorder: Protest and Politics in Italy, 1965-1975*. Oxford: Oxford University Press.

Taylor, Charles Lewis, and Michael C. Hudson. 1972. *World Handbook of Political and Social Indicators*. 2nd ed. New Haven: Yale University Press.

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

Tilly, Charles. 2008. *Contentious Performances*. Cambridge: Cambridge University Press.

Zhang, Han, and Jennifer Pan 2019. "CASM: A deep-learning approach for identifying collective action events with text and image data from social media." *Sociological Methodology, 49*(1), 1-57.

Torres, Michelle. Forthcoming. “A framework for the unsupervised and semi-supervised analysis of visual frames.” *Political Analysis*.

Online Appendix - Observation template: Dimensions coded in the analysis of Covid-19 protests in Italy and Germany

1. Meta Data

For all units, we collect the following meta data

- Link to the video
- Name of video file
- Day of the event
- Day of upload
- Who took the video
- Who published the video
- Duration of the video

2. Event

2.1. Organizers of the event

Which group(s) is organizing the event

2.2. Adherents to the event

Which other groups adhere to the event and are publicly present

2.3 Phase

Founding phase, storming the wealth day etc.

2.4 Duration of the event

2.5 Location of the event

City and location of within the city, e.g. square, municipality. Please provide a description of the location.

2.6 Form of the event

Demonstration, static sit-in, riots, etc.
Please also signal whether the event was allowed/prohibited

2.7 Presence of violence

Physical violence: damages to public properties, threats to journalists, etc.
Symbolic violence

3. Participants

3.1 Number of participants

Total number of participants.
As it is likely that the number will fluctuate, please provide the minimum and maximum number.

3.2 Has there been significant fluctuation in the number of attendants?

Records major fluctuation during the session.
Has there been significant fluctuation in the number of attendants? At the beginning/the end? Related to which issue? Note only if fluctuation exceeds 20 percent.
Major fluctuation can be a sign of "voting with the feet": participants may show their dislike of an issue/discussion by choosing the exit option.

3.3 Gender of participants
More male or female participants.
3.4 Age of participants
Age group of participants, e.g. younger (less than 30-35 years old) or older (above).
3.5 Presence of police
Is the police present at the event? From the beginning or later during the event? In which forms (visibility, equipment)?
3.6 Presence of counter-protestors
Are counter-protestors present during the event? From the beginning or later during the event? To which social or political group do they belong? What do they do?
3.7 Presence of bystanders
Presence of people that are not participating in the organization of the event (e.g. "spectators").
3.8 Atmosphere at event
General description of the overall atmosphere at the event. Major interruptions, applause, are people chanting? Which format for posters/signs etc.?

4. Interactions

4.1 Interactions among protesters
Friendly, quarrels, physical fights
4.2 Interactions with the police
How are protestors interacting with the police (e.g. peaceful, tension, escalation into violence etc.)
4.3 Interactions with counter-protestors
What is the interaction between protesters and counterprotesters, are they directly responding to each other's or are they kept separate by the police? Is there any escalation? Are there moments in which the interaction is more intense, which ones?
4.4 Interactions with bystanders
Are the interactions peaceful, friendly, or rather tense and aggressive?
4.5 Interactions with journalists
Are the interactions peaceful, friendly, or rather tense and aggressive?

5. Communication

5.1 Speakers
Please list all the speakers and describe who they are.
5.2 Length of speech
5.3 Slogans
Please provide a list of slogans
5.4 Posters, flags, symbols
Please provide a list of posters, flags, signs, symbols.
6. Frames
Within the transcripts, we code statements about the Corona pandemic. We are not only interested in statements that directly refer to the virus, but also in descriptions of problems, related statements on conspiracy theories or calls for action. Each occurrence of the statements below is coded individually.