

## 2 | *The Methods and Methodology of Micro-sociology*

This chapter introduces micro-sociological methodologies and analytical strategies, including the ontological and epistemological underpinnings of micro-sociology. The chapter proposes three analytical strategies for micro-sociological analysis: studying key events, interaction ritual chains, and patterned phenomena. Micro-sociological analysis can be conducted with various methods, from interviews and participatory observation to textual analysis, surveys, and digital methods. In particular, the chapter focuses on video data analysis (VDA), because this method is particularly well suited to capturing micro-dynamics of rhythm, emotions, and bodily interaction. The chapter shows how to gather, code, and analyze video material and illustrates how VDA can be triangulated with other methods. While VDA lends itself well to positivist studies coding, counting, and replicating observations, VDA can also be applied from a post-positivist approach and holds quasi-ethnographic potential. Whatever the epistemological standpoint, VDA is related to several dilemmas related to (1) access and availability, (2) validity and veracity, (3) data presentation, and (4) ethics, which will be discussed in this chapter. The chapter thus seeks to provide input to students and researchers interested in applying micro-sociology in studies of peace and conflict, not only in terms of how to conduct VDA and what to study but also the epistemological choices and potential problematics involved in doing so. Finally, the chapter presents the data sources, methods, and methodological considerations that make up the empirical basis of this book.<sup>1</sup>

<sup>1</sup> Elements of this chapter were previously published in *Conflict, Security & Development* by Taylor & Francis Group on September 27, 2022, available online: [www.tandfonline.com/doi/full/10.1080/14678802.2022.2122696](https://doi.org/10.1080/14678802.2022.2122696).

## The Ontology of Micro-sociology

Before proceeding to a presentation of micro-sociological methods, a note on the ontological underpinnings of the micro-sociological approach is in order. Collins (2004, 16) writes that “Goffman is a social constructivist, except that he sees individuals as having little or no leeway in what they must construct; the situation itself makes its demands that they feel impelled to follow.” In a similar way, one can say that, micro-sociology is a form of social constructivism, albeit not in the most common use of the word. Unlike constructivists, who perceive social life to be a product of ideas, norms, or discursive deliberation, the productive unit in micro-sociology is the interaction ritual and specific situations. Interaction rituals produce solidarity, emotional energy, symbols of social relationships, and standards of morality, which are what make up the pillars of a society. Thus, micro-sociology is social construction in its very basic form; symbols such as a national flag or moral conduct like the Danish “Law of Jante”<sup>2</sup> are not (just) constructed in the human mind but emerge from social situations and interactions. In other words, the micro-sociological approach is more social than constructivist. A more precise description might be *social emergence*, since it is the product of individuals’ interactions – not conscious ideas about what to construct – that produces social life. Whereas many social constructivist approaches assume that our recognitions and perceptions of the world produce (or rather: is) the (social) world (Collins 2012), the reverse is the case in the situational account, where our perceptions are largely seen as emerging from concrete situations (or, as Collins (2004, 345) adds, from interaction rituals within the mind). The fact that emotional energy, solidarity, and symbols of social relationships emerge from social interaction makes them no less “real.” The social world is shaped by certain mechanisms, not laws, that exist independently of our realization of them (in fact, many people often do not consciously recognize social mechanisms; they just *feel*, e.g., that something is wrong in tense situations, or they feel dispirited in dominated situations).

<sup>2</sup> The “Law of Jante” is a Danish code of conduct created by Aksel Sandemose, a Danish author, in a novel entitled *A Fugitive Crosses His Tracks* (*En flyktning krysser sitt spor*, 1933). It essentially states that nobody should believe that they are better than others, and is widely regarded as a key norm in Danish social life.

The situation is the core starting point for Collins' micro-sociological theory. In fact, according to Collins (2009a, 21), even epistemological and ontological problems are produced in concrete situations:

[T]he whole of human history is made up of situations. No one has ever been outside of a local situation; and all our views of the world, all our gathering of data, come from here. Philosophical problems of the reality of the world, of universal, of the other minds, of meaning, implicitly start with this situatedness.

Collins' theory builds on Goffman's methodological situationalism (Jacobsen 2012) but leaves greater room for agency with his theory of emotional energy. Emotional energy is a force of agency; individuals with little emotional energy will find it difficult to make decisions and act, whereas those with high emotional energy are able to set big events in motion and define the rhythm of the interaction rituals in which they take part. Emotional energy is generated in social situations but also persists a given amount of time thereafter, and it is therefore input in other social situations. Collins, therefore, speaks of chains of interaction rituals that feed into each other. In this way, micro-sociology presents a different take on the structure–agency question that remains material for eternal academic discussions (Demmers 2012). Structure is not an invisible force operating over and above micro-interactions; rather, it is an emergent phenomenon, composed of micro-interactions. Likewise, agency is not a given, fueled instead by emotional energy generated in micro-interactions.

What is the ontological status of emotional energy? Does it exist beyond the human mind? Collins would argue that it does, arguing further that the level of emotional energy can be measured as the relative difference in the hormone testosterone (not how much you have but how much you usually have, depending on whether you are female or male). Moreover, emotional energy is also observable in facial expressions and voice (Collins 2004, 133–9). What, then, is the ontological status of interaction rituals? Are they merely heuristic ways of explaining human interaction? Or are they also biologically wired? Several elements of interaction rituals can be explained biologically. In particular, the assumption about the human tendency to get rhythmically entrained in bodily copresence corresponds to neurobiological findings about human nervous systems becoming “mutually attuned” (Collins 2004, 64). Collins (2004, 78) therefore concludes that

“emotional contagion is a socio-physiological fact . . . From an evolutionary perspective, it is not surprising that human beings, like other animals, are neurologically wired to respond to each other.” This is supported by Heinskou and Liebst (2016), who further specify the neurobiological features of Collins’ interaction ritual.

That emotional energy and rhythmic entrainment correspond to biological tendencies does not per se lead to the assumption that human beings are motivated by striving to maximize emotional energy and engaging in intense interaction rituals. This assumption is an ontological statement equivalent to the rational choice assumption that human beings seek to maximize utility in any given situation. As mentioned in Chapter 1, this book does not support the idea that human beings are always guided by the aim to maximize emotional energy. Rather, I follow Pouliot (2008, 276), who argues that the logic of practicality is ontologically prior to ideas and rational choice, because “it is thanks to their practical sense that agents feel whether a given social context calls for instrumental rationality, norm compliance or communicative action.” In other words, it depends on the situation whether actors will follow, for example, rational calculations. In some situations (e.g., trade) rational calculations are appropriate, whereas other logics will be required in others (e.g., raising a child). Hence, it depends on the situation whether actors are (primarily) guided by rational calculations or other logics (Collins 2004, 141–81).

### **Micro-sociological Research Strategies**

Micro-sociological analysis can take many forms, depending on the research question and availability of data. Generally, at least three overall analytical strategies (Andersen et al. 2005) for micro-sociological analysis in peace research can be identified: studying key events, interaction ritual chains, and patterned phenomena.

#### *Key Events*

First, micro-sociological studies can imply the analysis of a particularly significant or rare event shaping international conflicts via video recordings of the event and/or thick descriptions from, for example, war memoirs, interviews, or diplomatic biographies. Several events of relevance for world politics are recorded, which allows for the most

fine-grained micro-sociological study via VDA. This allows the researcher to grasp how certain critical junctures unfold at the micro level by considering, say, how a particular political speech is constructed interactionally, emotionally, and affectively in the moment. Rather than (merely) analyzing the symbolic and rhetorical dimensions of these events, VDA opens up for the analysis of facial expressions, body language, and the interactions between the participants. For example, Klusemann (2009, 2010, 2012) studies the Srebrenica genocide (1995) by analyzing eight hours of video footage of the events recorded by a Serbian camera team. Klusemann conducts a moment-by-moment sequential analysis of the recording, coding verbal as well as nonverbal behavior and emotional cues based on, among other things, the methods for detecting emotions in facial expressions and body language developed by Ekman and Friesen (1978). When studying key events, it can make sense to study the rising and falling levels of intensity in the interaction, such as measuring the tempo in the rhythm of interaction in a demonstration, a diplomatic meeting, or an attack.

### *Interaction Ritual Chains*

Besides analyzing specific events, micro-sociological studies can also investigate chains of interaction rituals that together form or lead to a particular world political event. Pouliot (2015) coins his approach to process tracing “practice tracing,” indicating that the unit of analysis as well as the force that is believed to bring matters forward is practices. A Collins-inspired analysis could thus be said to conduct *interaction rituals tracing* (Bramsen 2017, 55). Here, tracing chains of interaction rituals, researchers can investigate how actors are energized or de-energized and connected or disconnected in certain situations and how this feeds into new interaction rituals. It is difficult – in many cases impossible – to collect all of the micro-situations comprising a given macro-social phenomena. Instead, Collins (1983, 194) argues that researchers can investigate a sample of representative situations and “fill in the rest by extrapolation.” In my study of nonviolent and violent trajectories of conflict, for example, I have traced and compared micro-dynamics of the unfolding of events in the initial phases of the Arab Uprisings in Bahrain, Tunisia, and Syria (Bramsen 2017). I use videos of demonstrations in the respective countries to understand how the movements were energized by engaging in protest activities

and whether security forces were able to dominate protesters. I couple this with interviews with activists, journalists, and opposition politicians as well as news media and human rights reports to obtain a picture of how concrete interactions between activists and security forces in the streets shaped the overall power balance and unfolding of events. The result is a granular depiction of the “chains” through which the Arab Uprisings developed through micro-events “on the street.”

### *Patterned Phenomena*

Finally, another option is to trace patterns across categories of events – globally or locally – to provide a broader picture of the micro-dynamics of world political events. Here, the focus is not on how different interactions feed into each other and comprise larger developments; rather the focus is on what characterizes these types of interaction, such as acts of torture, demonstrations, or diplomatic meetings. For example, Austin (2020) investigates the “just-whatness” of violence in over 200 videos of torture, primarily from Syria, and finds that while tools for torture are circulated globally, violence is a locally performed practice shaped by negotiation gestures, rhythmic entrainment, and practices to keep the violence moving, all of which points to the difficulty of conducting violence. Likewise, Anisin and Musil (2021) analyze 147 videos from the Gezi Park uprisings in Turkey and show how attempts made by the protesters at fraternizing with the police often affect the interaction. The three approaches and their focus, data, and potential problems are illustrated in Table 2.1.

### **Video Data Analysis**

Various methods can be applied in a micro-sociological study, including analysis of war memoirs (Mac Ginty 2021, 2022a) and biographies (Holmes and Wheeler 2020). One of the most fitting methods for micro-sociological research is VDA as either primary or supplementary method, as it enables researchers and students to grasp and analyze fine-grained details of interaction, such as facial expressions, body language, tone of voice, and rhythm of interaction. Peace research has always been open to new methodologies and theoretical approaches capable of shedding light on previously overseen causes

**Table 2.1** *Three micro-sociological research strategies*

	Key events	Sequences	Patterned phenomena
<i>Focus</i>	Critical events that are widely agreed to have had significant world political consequences.	An analysis of how the dynamics of particular events evolve over time, in ways that are linked to antecedent events, due to fluctuating variables resulting in a particular eventual outcome.	The identification of more or less generalizable patterns across discrete phenomena of the same type.
<i>Data</i>	Video material of the event concerned, interviews, participant observation, transcripts of, e.g., diplomatic meetings, reports.	Multiple videos of different linked events across a temporal period leading up to a particular outcome, interviews, participant observation, reports.	A sample of videos representing the patterned phenomenon in question (not necessarily linked temporally or spatially), reports, interviews.
<i>Problems</i>	Designation of an event as “key” is often caused by mainstream media or history writing that may undermine marginalized voices and alternative interpretations.	Important events in sequences may not be video recorded.	Unequal availability of video data.
<i>Examples</i>	Ariel Sharon visits Temple Mount; Declarations of Victory.	Small-scale peaceful protests beginning at time X that eventually attract a violent police response at time Y.	Any phenomenon that occurs regularly but whose occurrences are not necessarily specifically temporally linked (torture, protests, speeches, negotiations, etc.)

and dynamics of peace and conflict (Wallensteen 2011a, 17). Likewise, this book pushes the methodological boundaries for peace research by introducing VDA as a new method in peace research. Thus, the book furthers the aim of “making peace researchable” (Wallensteen 2021) – by not merely studying declines in battle death nor the structural conditions for peace, but also by examining the relational dimension of how people *interact*.

Video data analysis has been applied in sociology and psychology for many years (Kjær and Davidsen 2018), including conversation analysis, but has more recently been adopted as a method for more general use (Nassauer and Legewie 2018, 2022). An increasing number of qualitative and quantitative scholars alike have begun employing visual data, including photographs, motion pictures, video clips distributed on social media, and artistic representations (Bleiker and Butler 2016). Unlike visual International Relations (IR), VDA is not focused on the aesthetics of politics (Bleiker 2009) or political effects of specific images (Hansen 2015). Rather, videos are analyzed to understand the interaction they portray. Videos are also not used to *document* or *prove* real-life events; rather, they serve an observational purpose, providing a window through which to observe real-life events, including the atmosphere, sounds, rhythm of interaction, body language, facial expressions, and contextual factors. In this way, visual data offer analytical potential that is complementary to participant observation (Nassauer and Legewie 2018).

Researchers are rarely at the right place at the right time, such as when violence occurs or is deliberately prevented through nonviolent gestures. Here, video data come in handy. Surprisingly, many things are recorded and available online; from Trump’s dominating handshakes with other heads of state to fighting on the ground in Syria. In many cases, video material allows researchers to go *back* or *far away* and observe events at the right time and place, because people who happened to be there recorded it or photographers recorded events for the news media. In this manner, videos enable the researcher to observe events from their armchair to which no (or few) researchers would otherwise have access, integrating some of the detail and attentiveness to interaction and atmosphere that only ethnographers would have. While losing the ethnographer’s benefit of being present and able to see the whole scene to a greater extent, videos allow researchers to replay events in slow motion, thereby capturing subtle, hardly noticeable



dynamics, such as changes in tone of voice, pauses in speech, or how participants mirror each other. Video data are therefore often a source of surprise that can challenge traditional understandings of a phenomenon, inspire a reconceptualization of theory (Bramsen and Austin 2022), and “generate completely new insights” about social life, peace, and conflict (Nassauer and Legewie 2022, 5).

Adding to this, videos may almost serve an ethnographic purpose of giving researchers a sense of being present in particular events. Although video material neither transmits smell nor provides the opportunity to engage with the people present on the scene, it does provide not only visual but also auditory insights into a given setting. In my own research, for example, I have used video material from protests in the Arab Uprisings to analyze micro-dynamics as well as to better understand the atmosphere and participant perspectives:

Watching hundreds and hundreds of videos of people chanting rhythmically in demonstrations combined with interviews with informants’ graphic recounting of the events occasionally gave me a sense of almost “being there” – a historic window into the Arab Uprisings provided not only through words but also sounds and visuals, which often left me with revolutionary songs stuck in my head long after watching the videos. (Bramsen 2017b)

Hence, video material holds quasi-ethnographic potential. But unlike participant observation, video data do not just give us “one-shot” at accessing world political events but have the advantage that the observed incident can be replayed repeatedly and thus analyzed in micro-detail that is rarely captured with ethnographic methods, including the intonation of speech, facial expressions, body language, and the rhythm of interaction (Collins 2004; Liberman 2013).

Some questions are more prone to micro-sociological analysis than others, both due to the availability of data and the mechanisms at play. Social movements and mobilizations are particularly prone to micro-sociological analysis, as such cases lend themselves to the analysis of what gives people energy to go to the streets or take up arms, how momentum for uprisings spreads, and who ends up with the upper hand in a struggle (Bramsen 2017; Solomon and Steele 2017). In modern times, where most protests are filmed, these cases are also relatively easily observed and analyzed via VDA. While it is more difficult to obtain video data from other aspects of peace and conflict, such as elite negotiations, peacebuilding activities, or trade wars, it is

possible to use other data such as surveys, interviews, or simply to record relevant activities yourself.

## **How to Conduct VDA**

A VDA study essentially consists of a three-step process: (1) data collection, (2) coding, and (3) analysis. Each of these steps is outlined in the following.

### *Data Collection*

The first step of VDA is the collection or recording of videos. Videos can be recorded by researchers themselves, found in documentaries, obtained via access to, for example, CCTV, or collected online. First, the benefits of recording the video as a researcher is that it then becomes an add-on to ethnographic studies where the researcher had both the benefits of experiencing a given situation and being able to interact with people there while at the same time being able to replay key aspects of the observations and thus observe them in micro-detail. This also gives the researcher access to situations that are otherwise not recorded at length. Lund (2017), for example, has recorded and analyzed peacebuilding activities in Uganda. If the researcher records the video, they must obtain consent from all the actors figuring in the video.

Second, video material can also be found in documentaries. Here, it is critical that the researcher is aware of any manipulation possibly carried out by the producer of the documentary, and it is a good idea to triangulate the video material with other sources, such as interviews with participants. Also, if possible, it is best to obtain all the raw data from the documentary. For this book, I use two documentaries, for one of which I have access to the raw data. In both cases, I also interviewed the mediator facilitating the dialogue in the documentary so as to get insights into the experienced situation, the role of the camera in potentially shaping the interaction, and the details not captured on video.

Third, several VDA studies have also applied CCTV (e.g., Philpot et al. 2020) or other recordings not available through open access but through application for access. For example, the International Criminal Tribunal for the former Yugoslavia gave Klusemann (2010) access to eight hours of video footage that had been used during trials.

This video was recorded by a Serbian camera team following the events leading up to the Srebrenica massacre.

Finally, video data can be collected open source. Smartphones and photojournalists are increasingly recording aspects of social life for non-research purposes. This material is highly useful for researchers and potentially provides insights into everyday situations or iconic events of relevance for world politics. Videos can be found on Google, YouTube, and Vimeo, as well as on social media platforms, such as Facebook, TikTok, and Twitter. When I analyzed the Arab Uprisings in Bahrain, Tunisia, and Syria, for example, I found the protester Facebook groups and searched back in time to get the visual material posted early in the uprisings. Likewise, when I analyzed the Philippine peace process, I was able to access most of the video and photo material from the Facebook pages of the government and the communist party. The collection of videos often involves an element of trial and error, as search terms are varied, multiple languages are (ideally) employed, and different types of outlet are searched (e.g., mainstream video platforms and social media platforms). Beyond online platforms, interviewees may be helpful in pointing toward relevant videos, or interviewees themselves may have recorded events relevant for the research.

### *Coding*

Following the data collection process, the video data can be coded to unpack the contents of a set of videos in terms of the observed body language, the nature of the interactions visible, the types of material artifacts involved, and/or the use of language. The coding may be strict and produce numbers that become central findings in the research, but it may also be done solely for the researcher to systematize the data and to find patterns that characterize the videos. Coding may be conducted manually or automatically.

Several methods exist for the automatic coding of video material, coding facial expressions, sounds, actions, or recognizing objects. One group of tools for automated coding is applications with graphical user interfaces (GUI), where a simple click and drag with a mouse is enough to classify and count specific visual patterns of a number of videos. The Noldus Face Reader, for example, offers automated face recognition and the coding of emotions. Another group of tools consists of code

libraries that can analyze video material through the environment of statistical programming languages, such as R or Python. This group offers greater flexibility in terms of what and how patterns should be classified and counted, although they must be developed and trained for each new research project and is, hence, very costly. Finally, a group of so-called Cloud AI has emerged in recent years. Cloud AI are services offered and developed by huge tech companies (e.g., Google, Microsoft) offering researchers and analysts easy ways of mobilizing artificial intelligence to analyze video material. Researchers can upload their material, for example to a Google server, and receive analytical results provided by the highly advanced image and video algorithm Google has developed over time for carrying out searches on YouTube, Google, and other platforms. At least theoretically, the Cloud AI thus offers the same capabilities as the code libraries but with the ease of the graphical applications. It remains too early to determine if Cloud AI will be able to deliver on its promises made within the area of automated video analysis.

### *Analysis*

When analyzing (and manually coding) video data, the data can be watched repeatedly (and in slow motion) to obtain a detailed picture of how actions evolve. This opens up whole new possibilities for analyzing situational details, counting speaking time, objects, or instances – something that would not be possible with ethnographic observations. In some cases, it can be helpful to watch the video without sound to be able to focus on other important details. As in any analysis, a crucial element of VDA is looking for patterns in the data of analytical relevance (Nassauer and Legewie 2019). Patterns may be found by counting specific things, looking for changes or turning points in interaction, tracing temporal development, or ordering practices. Often, it is beneficial to look for surprising elements in the video data; something that challenges or eloquently exemplifies established understandings of war and peace: What stands out as different from what a theory would predict? What surprised you the most when watching the videos? Is there an interesting pattern across different videos? In my study of interaction in the Northern Ireland Assembly (Chapter 5), for example, I noticed how politicians refused to clap when opponents were elected even though they were the ones promoting their candidacy in the first

place, which says something about the theatrical nature of the oppositional interaction in the assembly (Bramsen 2022a).

### **Methodological Triangulation**

When analyzing complex phenomena and sensitive events that are difficult to observe directly, multiple methods become crucial. VDA is essentially situation-bound (Nassauer and Legewie 2022, 39), often not revealing what occurred immediately before and after the video was recorded (e.g., what happened in the corridors immediately before a diplomatic meeting). In many instances, VDA can therefore benefit from being coupled with or supplemented by other methods. In particular, methods of ethnographic observation, deep textual analysis, ethnographic interviewing (Spradley 1979), discourse analysis, content analysis, and more are all likely to be useful in overcoming “incomplete” visual information, uncertainties about veracity, and so forth. Triangulating VDA with interviews and/or participant observation can add to a study with insights into (1) the cultural and social contexts, (2) participant’s own experiences, and (3) what happens when the cameras are turned off or in the blind spots.

Whereas positivistic approaches to VDA emphasize methodological triangulation as a means to ensure a more complete capture of an event (synthesizing different data sources, such as police reports, news reports, interviews, videos, and court data, such as reconstructing a demonstration from a to z) and thus validation of the study (Nassauer and Legewie 2018), triangulation can also be used to provide a fuller picture of a given phenomenon without the different items of data necessarily coming together to reconstruct all of the aspects of an event. Different data sources need not provide different entry points to comprehend the *same* violent events but may also simply multiply the number of the situations that can be analyzed and add a *different* dimension to the study; for example, by taking into account the experience of conducting and being subjected to violence. Rather than ensuring complete capture, triangulation can thus provide a more comprehensive, overall picture of a given phenomenon.

One of the most obvious methods to be coupled with VDA is interviewing. A more in-depth understanding of a given video can be achieved by interviewing the participants in the video about their experience of the situation, their perceptions of the atmosphere, and

the more visceral dimensions of the recorded situation as well as their experience of what happened before and after the recording. Even if it is impossible to interview the exact participants in the available video material, it can be valuable to interview participants experiencing similar situations.

Unlike most qualitative, in-depth interviewing, the main objective of micro-sociological interviews is neither to obtain insight into the interviewee's lifeworld nor to understand how they give meaning to particular phenomena, but rather to investigate specific situational dynamics. Describing a situation in detail often does not come naturally to informants. The efficacy of interviews in IR is frequently limited by how respondents often simply repeat reflexively constructed autobiographical narratives of their lives rather than articulating the contours of their actual experiences in practical or ethnographic detail<sup>3</sup> (Bramsen and Austin 2022). The interview technique applied in micro-sociological analysis therefore implies questioning specific details and continuously nudging informants away from describing overall narratives of a phenomenon or an event to instead explaining micro-details and specific situations, body language, material artifacts, emotions, and interactions (Collins 2012, 3).

### **Epistemologies of VDA**

The key question in epistemology is what researchers are able to say about social life. The sociological approach to VDA has thus far primarily been generally positivistic, posing the complete capture of "natural behavior" as a criterion for validity. However, VDA is not inherently positivistic and might also serve more reflexive strands of research (Bramsen and Austin 2022). As we shall see in the following, the main bone of contention between the two approaches is the degree to which VDA provides access to how actors *actually* behave and interact or whether such "natural behavior" actually exists and can be captured on camera.

<sup>3</sup> Here, one should note that while this tendency of interview respondents may be deliberate, it is not usually intentional; instead, it reflects a general human tendency to fit the events of our lives within an autobiographical narrative that allows us to make sense of everything that has occurred therein while paying little attention to situational details (see Damasio 2012).

### *Positivist VDA*

In the Nassauer and Legewie (2018, 2019) sociological introduction to VDA, they propose a positivist and behaviorist approach to studying video recordings. The goal in the positivist approach to VDA is often to identify causal links. Similar to qualitative process tracing methods, this often implies “reconstructing a situation step-by-step” so as to “analyze its inner dynamics, and establish comprehensive story lines” (Nassauer and Legewie 2018, 15). For example, Lindegaard et al. (2017, 1) establish that “consolation in the aftermath of robberies resembles post-aggression consolation in chimpanzees” based on CCTV footage of robberies.

In the positivist approach, video data captures *natural behavior*, defined as the degree to which “actors in visual data behave in a way they normally do in the type of situation under investigation” (Nassauer and Legewie 2018, 23). Locating data meeting this standard is considered a key criterion for the “validity” of any study of visual data. Another criterion for validity within a positivist approach to VDA is “complete capture,” meaning using videos that portray a given event, the object of study, from one end to the other (Nassauer and Legewie 2018). Should this be unobtainable, methodological triangulation can be applied to ensure complete capture. For example, Nassauer (2019) puts together different data sources (e.g., police reports and media articles) to compose a full picture of what happened second-by-second in demonstrations in the United States and Germany.

The positivist, behavioristic epistemology intuitively fits the VDA method, as it enables the study of behavior and, unlike ethnography and interviews, the researcher is not shaping the results by virtue of being in the same space as the informants, instead able to observe them from afar, possibly even hiring research assistants to code the video material. Likewise, the positivist ideal of replicability is possible when applying VDA as long as the analyzed videos are publicly available – or at least available for other researchers. However, it is also possible to apply VDA within other epistemological frameworks, as we shall see in the following.

### *Post-Positivist VDA*

While studying behavior using VDA may be intuitively linked to behavioristic analysis, it is not innately “positivistic” in analyzing the

“behavior” or “practices” (Pouliot 2008) of human beings, as they are depicted in visual artifacts as long as the analyst avoids an *a priori* search for “natural behavior.” Students of IR who take a more “subjective,” “interpretivist,” “reflexive,” or – simply – “critical” approach to exploring world politics should therefore avoid the temptation to read positivistic sociological variants of VDA as anything other than *one* deeply contested understanding of how VDA can proceed. That said, the methodological difficulties faced by scholars who refuse such a positivist behavioral reading of visual data are significant. The multiple layers of meaning that visual artifacts are infused with, as well as the multiple possible ways in which one can “read” the depiction of events, practices, or situations (even in behavioral terms), means that the analysis of visual artifacts must inevitably constitute an iterative process, the conclusions of which can only ever be tentative and contingent.

One of the concepts in the Nassauer and Legewie approach that grates on the ears of post-positivist scholars is “natural behavior.” While distinguishing between “staged” or “un-staged” behavior (e.g., differentiating between violence depicted in Hollywood films portraying World War II and videos filmed during World War II itself) clearly makes some sense, it remains deeply problematic from a post-positivist perspective to consider natural behavior as something that is simply “out there” to be captured and which can be studied independently from the observation of the researcher or even the person recording the video. To some degree, assumptions that visual data can be more “objective” than other forms of data stem from the false belief that any camera recording this data can operate as a “neutral” observer, whereas, say, an ethnographic observer embedded in a particular situation cannot. This obviously misses the degree to which the camera itself not only influences the event occurring but also how – with Butler (2009, 66) – the camera positioning is a way of interpreting in advance: “[A]lthough restricting how or what we see is not exactly the same as dictating a storyline, it is a way of interpreting in advance what will and will not be included in the field of perception” (Butler 2009, 66).

In Nassauer’s (2019) empirical study of street demonstrations, she argues that because demonstrations are often recorded by police, demonstrators, and the media, the actions in a demonstration constitute natural behavior, as it is considered quite normal to have actions recorded in these circumstances and – hence – the behavior of



individuals remains somehow “natural.”<sup>4</sup> However, one might argue that this permanent presence of recording mediums during events like protests actually *demand*s a deeper accounting of their role in influencing behavior: What would demonstrations look like without cameras? To what degree are protesters and/or police “acting” for the cameras? To some degree, how might key emotional, affective, and/or discursive aspects of demonstrations be *missed* by these recordings? And so forth. Such questions indicate that analyzing situations, events, or sites where cameras *do* affect social interactions should not be considered an “invalid” research practice. On the contrary, this fact demands only deeper inquiry into the multiple possible layers through which any image can be interpreted. Indeed, in many cases, the presence of cameras should not be treated as a “potential bias” to be taken into account, but rather as an inherent aspect of the interaction. This is especially so in diplomacy, where a performance like handshaking is conducted precisely for the benefit of the watching cameras. In Chapter 6, for example, I analyze how the presence of cameras at press conferences seems to energize participants in the Philippine peace talks, and in Chapter 7, I discuss the performativity of diplomacy.

In a positivist application of VDA, researchers need not necessarily have any greater in-depth knowledge of the cultural context in which a given video is recorded as long as the lack of cultural knowledge does not disable the researcher to catch small cues or cultural variations (e.g., in smiling).<sup>5</sup> In the post-positivist tradition, context-specific knowledge and understanding are essential, and methodological triangulation is therefore not applied to obtain the “full picture” of a particular situation but rather to get more in-depth understanding of the case and context. For example, interviews and visiting the places studied (including the specific sceneries, like the location of a protest) add to the researcher’s contextual and cultural embeddedness.

<sup>4</sup> Collins’ approach to analyzing visual data is likewise behavioristic, although his ontological standpoint is different than that of most behaviorists, arguing that behavior is shaped by situational circumstances and generated in human interaction, not a product of simple stimuli–response mechanisms.

<sup>5</sup> According to Ekman and Friesen (1978), core emotions of anger, disgust, fear, joy, sorrow, and surprise are expressed relatively similarly in facial expressions despite cultural differences.

## Dilemmas of VDA

As with any method, there is a set of challenges and dilemmas inherent in employing VDA, particularly in IR and peace research. In this section, four specific dilemmas intrinsic to VDA will be discussed; difficulties relating to the access to visual artifacts, validity, data presentation and ethics. While each of these dilemmas can render visual analysis a complicated method, none are insurmountable if careful consideration is given to how they affect the design of research using visual methods of analysis.

### *Access*

While an ever-increasing numbers of events and interactions relevant to peace research are now recorded and made available, many other practices, situations, and events of relevance remain that are unavailable in visual form. This is typically because these phenomena are confidential, private, or simply not recorded for various reasons. For example, it might be highly useful to analyze the micro-dynamics of President Bashar al-Assad's interactions with his family and advisors during the initial phase of the Syrian uprising or to directly observe the peace talks in Colombia. For good reason, however, these are not visually documented and therefore unavailable for analysis. Nonetheless, visual artifacts of many secretive, confidential, and/or controversial practices are increasingly made available. This even includes the leaking of videos of war crimes, including more hidden violence, like torture. Such access will possibly continue to be increasingly provided through leaks, happenstance, and/or releases through freedom of information requests in the future.

A word of caution is necessary here, however. It is possible that the foci of VDA risks being driven more by data availability than relevant research questions focused on relevant material. One particular issue here is likely to be the much greater quantity of visual data available from non-Euro-American states depicting practices of violence, abuse, corruption, etc. This material typically becomes more available in less wealthy states due to the lesser resources available to these authorities to control the release of data by personnel, foreign governments, or even hackers. The issue here is that depictions of, say, war crimes by the Syrian government become "hyper-visible," both publicly and

within VDA, whereas the war crimes of, say, the United States in Iraq, Afghanistan, and elsewhere are rarely released in visual form. The risk then becomes, that particular sociopolitical binaries depicting the Euro-American world as more “civilized” than states elsewhere are falsely reinforced. That said, VDA is not very different from other methodological approaches in this respect. Rational choice analysis, for example, does not have full access to the calculations of political leaders and their followers and must rely on assumptions and proxies. Moreover, as described earlier, VDA may very well be supplemented and “triangulated” by other methods, like interviewing and participatory observation (or even discourse analysis or quantitative studies) to pursue the relevant research inquiries to mitigate these problems.

### *Validity*

Assessing the “validity” of visual artifacts gathered for the purposes of VDA involves many difficulties. These difficulties are not unique to social scientific analysis. For example, Wessels (2016) has demonstrated how only a fraction of YouTube videos depicting violations of human rights or war crimes in Syria can be used for legal evidence in future prosecutions (of whatever type) due to the lack of verifiable contextual information indicating the date and time the video was recorded, the geographical location in which the events depicted occurred, and – most importantly – the identities of the perpetrator and victim. In addition, the very “truthfulness” of visual artifacts is often contested, with numerous “fake” videos frequently appearing that feature staged events. Likewise, it is difficult to assess the validity of videos for social scientific purposes.

For some, the ideal visual artifact for social scientific analysis seeking to achieve a “comprehensive” overview of a particular phenomenon would be an artifact capturing the entirety of a particular situation, event, or phenomenon (e.g., popular protest, battle sequence, diplomatic negotiation). Preferably, multiple videos shot from several angles, in a high-resolution format, and with appropriate metadata would also be available. This set of requirements constitutes another set of “criteria for validity” suggested by scholars using VDA within sociology (Legewie and Nassauer 2018, 154–9). Nonetheless, I suggest a broader view here. Such comprehensive visual data is likely only to be available for phenomena captured, for example, through

high-definition CCTV cameras. Restricting our analysis to such cases would radically reduce the scope of VDA and its potential range of contributions. Moreover, it would also prevent VDA from appreciating the importance of, say, visual artifacts recorded on mobile phones.

However, what is likely to be more important than validity within VDA concerns the “veracity” of the images used in any analysis. Visual artifacts are increasingly being manipulated through various means (e.g., Photoshop, video editing software) or even “staged” outright for particular sociopolitical purposes. This has always, of course, been a problem vis-à-vis studying the visual, but the problem has only increased lately as its use has been embraced by governments and non-stop groups alike as a means to further their sociopolitical goals. More prosaically, “real” videos may be (deliberately or not) mislabeled, mis-categorized, or mis-located during their dissemination for various reasons. While it may not be possible to rule out the fakeness of a video completely, there are several ways to minimize this risk, including contextual knowledge, multiple videos of the same event (perhaps from different angles), data triangulation, and interviews with participants who can confirm or deny veracity.

### *Data Presentation*

One of the more pragmatic challenges in VDA relates to the presentation of research findings. It will always be difficult to communicate the contents of visual artifacts (and particularly videos) to readers through the textual mediums that still dominate social scientific meanings of disseminating research. The obvious solution for publishing visual data analysis here would be an expansion of the publishing model within IR to allow the easier linking of the textual versions of articles to visual materials, which would allow a reader to follow the analysis undertaken directly. In the absence of this possibility so far, the option for researchers is to present screenshots from a video within texts. Here, it can be valuable to construct small sequences of screenshots to illustrate the development in an interaction. However, there are several pitfalls to this method. First, some of the videos may not be of good enough quality for screenshots. While body postures and movements can be analyzed in a video despite poor camera resolution (e.g., from the initial stage of the Syrian uprising), screenshots are often more or less unreadable, especially when involving rapid body movements. Second,

screenshots might fail to capture important elements that are otherwise integrated into the analysis, particularly the centrality of sounds, interactions, and movement to VDA. Third, many journals would require the consent of the individual recording the video or even from all of those figuring in the photo/screenshot before publishing the article. This can be highly challenging to obtain, if for example recorded by protesters unwilling to reveal their identity or if they have ended up in prison or even been killed by the police. One way of circumventing the ethical challenges of printing screenshots is to instead have an artist edit or otherwise illustrate the photos in anonymous ways that nonetheless convey the facial expressions and body language necessary for the reader to visualize the situation. In their study of robberies, Philpot et al. (2020) have the photographs they use redrawn for the article. And with artificial intelligence improving by the day, it could conceivably become possible to apply the program Dall-E to re-visualize and anonymize photos or screenshots for published materials<sup>6</sup> (Solheim 2022). The issues related to data presentation are, of course, not unique to VDA. In fact, they are central to all forms of research (e.g., ethnography, narrative approaches, aesthetic approaches) that employ more or less unusual forms of data in their analysis.

### *Ethical Considerations*

Finally, applying video data implies a set of ethical considerations. The primary issue of concern revolves around the consent and safety of the participants in the analyzed videos. If recorded by a researcher, they naturally must obtain consent from the individuals featured in the video. If it is found online or is publicly available in other ways, it can be extremely challenging merely finding out who the participants in a video are, let alone obtaining their consent. There is no easy solution to this challenge, and in-depth ethical considerations must therefore be made prior to any VDA study to consider any potential risks or dangers for applying the videos. Here, whether the video is already publicly available is obviously important, as the added risks of using it in research may be limited. If further exposure of people can be

<sup>6</sup> This new technology obviously also poses a whole new set of ethical considerations in VDA: How can we know that a video is real and not fabricated or manipulated on a computer?

considered a risk to their safety, another option is analyzing the data while blurring faces (Nassauer and Legewie 2018). Even if the safety of an individual figuring in a video is not jeopardized by a screenshot from a video figuring in academic work, they may still wish to refrain from being an object of study. When people are part of online videos, they do not necessarily expect researchers to be observing them, especially if posted in a group of a more closed nature (e.g., on Facebook). While this would also apply to people in the public space, such as pedestrians on the street, many ethical guidelines nevertheless require consent from actors figuring in a video, of course particularly if screenshots are used for data presentation, as described above. Here, the main question facing researchers is: “Is this an invasion of privacy? And if so, to what extent?”

While the problem of consent can put limitations on what video material can be studied with VDA, it must importantly also be held against the social benefits of the study (Nassauer and Legewie 2022). As in any study, the essence of ethical considerations regards the pros and cons of analyzing video data; that is, the risks involved vis-à-vis the potential contributions of the study to society or the context under investigation. For example, analyzing protest videos would involve consideration of the protester perspective on research applications of the video material they have uploaded to the Internet. A crucial element of several ethical standards, including GDPR, is to avoid processing personal data. Anisin and Musil (2021) therefore deliberately removed any personal data in their data set of videos from the Gezi protests in 2013: “Whenever we noticed that a video includes personal details (e.g., the name of a protester or police officer), we eliminated it from our collection of data.”

## **Data and Methods in this Book**

This book draws on a number of different case studies, data sources, and methods. The different cases include the UN Security Council, international meetings between various heads of state, the Philippine peace talks, EU-led talks between Kosovo and Serbia, dialogue sessions between Kosovan and Serbian youth, Colombian peace talks and National Dialogue, the Northern Ireland Assembly, and the Arab Uprisings in Bahrain, Tunisia, and Syria. An overview of the different data sources and methods can be found in Table 2.2.

**Table 2.2** *Data, methods, and case studies in the book*

	Arab uprisings in Bahrain, Tunisia and Syria	Philippine peace talks	International meetings and peace talks	Kosovo-Serbia dialogue and negotiations	Dialogue and conflict transformation
<i>Data</i>	77 videos, 52 interviews, 5 human rights reports, participatory observation in a demonstration	5 videos, 5 interviews, participatory observations	8 videos, 30 interviews, participatory observation of NWM meetings and UN general assembly meeting	2 documentaries (one with all raw material available), 2 interviews	12 interviews, 12 hours of video from the National Dialogue, 1 hour of video from dialogue in Israel Palestine, 2 hour video of Northern Ireland Assembly meeting
<i>Methods</i>	VDA, interviewing, participatory observation	VDA, participatory observation, interviewing	VDA, interviewing, participatory observation	VDA, interviewing	VDA and interviewing

In total, the book thus builds on the analysis of 97 videos (approximately 25 hours in total), 70 photographs, 103 interviews, and participatory observation of a demonstration, a meeting in the UN General Assembly, two dialogue sessions, six meetings in Nordic Women Mediators (NWM), and one week of peace talks between the Philippine government and the CPP. The videos were coded manually and analyzed according to various elements, from the interactional dynamic of violence to dominant gestures and speaking time. The videos on violence in Bahrain, Tunisia, and Syria have been uploaded in chronological order to the webpage: <https://violence.ogtal.dk/>. Two of the videos that I draw upon are documentaries (*Reunion: Ten Years after the War* and *The Agreement*, the latter for which I have the raw material for the whole film. In both cases, I have interviewed the mediator in the documentary to ensure that I have as accurate an understanding of the interaction as possible and to understand what it meant for the process to have a camera in the room.

The fieldwork in Bahrain, Tunisia, and along the Syrian border was conducted in 2015 and 2016. In Tunisia, I stayed in Tunis for five weeks to carry out interviews with activists and to travel to the south, where the uprising started, where I stayed in locals' homes and interviewed activists. In Bahrain, I was unfortunately only able to remain one and a half weeks before being deported from the country on the grounds that my father-in-law (who was traveling with me to help care for my daughter) took a picture of a roadblock. In Bahrain, I interviewed activists, citizen journalists, and opposition politicians, and I participated in a demonstration together with the women in the back of the crowd. While I did not visit Syria due to the civil war, I did go to the Syria–Turkey border in Gaciantep, where many Syrian refugees stay, and was able to interview activists and citizen journalists.

The fieldwork investigating the Philippine peace talks was facilitated by my contact to Elisabeth Slättum, the special envoy to the Philippines at the time, and thus the chief mediator/facilitator of the peace talks. I was allowed to observe the third round of talks taking place in January 2017 at a hotel in Rome, where I also stayed. I signed a non-disclosure agreement promising that I would not publish anything from the negotiations until eighteen months after. To avoid interfering in the process, I did not conduct any interviews during my stay, but in 2020 I was able to maintain contact to the back-channel talks between the parties (taking place in Utrecht in the Netherlands), where



I conducted a number of interviews with the conflicting parties. Here, I also sat in on a pre-meeting between the Norwegian mediators and the communists (the CPP).

The fieldwork on the NWM meetings has been conducted together with Anine Hagemann. Since 2016, we have taken part in (and in several cases helped arrange) meetings in NWM. It was also in relation to a meeting with female mediators that I participated in a UN General Assembly meeting in 2017.

## **Conclusion**

This chapter has introduced the micro-sociological methodology in peace research, with specific focus on VDA. With more and more aspects of peace and conflict being recorded in high resolution, there is great potential for peace research to take advantage of these new data sources. The chapter has unfolded the ontological and epistemological underpinnings of the micro-sociological approach, recognizing the potential of both positivist and post-positivist approaches. Moreover, the chapter has sketched three analytical strategies one might apply within a micro-sociological framework: analyzing a specific and significant event, analyzing interaction ritual chains, and analyzing patterned phenomena. Like any method, VDA is not without dilemmas and pitfalls. The chapter has spelled out four central dilemmas of applying VDA to peace research; challenges related to access and availability of data, issues related to validity and veracity, challenges related to data presentation, as well as ethical issues. Finally, the chapter provided a short overview of the data and methods applied in this book.