

Nassauer, Anne; Legewie, Nicolas M.

Article — Published Version

Analyzing 21st Century Video Data on Situational Dynamics—Issues and Challenges in Video Data Analysis

Social Sciences

Provided in Cooperation with:

German Institute for Economic Research (DIW Berlin)

Suggested Citation: Nassauer, Anne; Legewie, Nicolas M. (2019) : Analyzing 21st Century Video Data on Situational Dynamics—Issues and Challenges in Video Data Analysis, Social Sciences, ISSN 2076-0760, MPDI, Basel, Vol. 8, Iss. 3, <https://doi.org/10.3390/socsci8030100>

This Version is available at:

<https://hdl.handle.net/10419/233033>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<https://creativecommons.org/licenses/by/4.0/>



Article

Analyzing 21st Century Video Data on Situational Dynamics—Issues and Challenges in Video Data Analysis

Anne Nassauer ^{1,*} and Nicolas M. Legewie ²

¹ Department of Sociology, John F. Kennedy Institute, Freie Universität Berlin, Lansstrasse 7-9, 14195 Berlin, Germany

² The Socio-Economic Panel at German Institute for Economic Research (DIW Berlin), Mohrenstr. 58, 10117 Berlin, Germany; nlegewie@diw.de

* Correspondence: anne.nassauer@fu-berlin.de; Tel: +49-30-83871283

Received: 8 February 2019; Accepted: 11 March 2019; Published: 19 March 2019



Abstract: Since the turn of the millennium researchers have access to an ever-increasing pool of novel types of video recordings. People use camcorders, mobile phone cameras, and even drones to film and photograph social life, and many public spaces are under video surveillance. More and more sociologists, psychologists, education researchers, and criminologists rely on such visuals to observe and analyze social life as it happens. Based on qualitative or quantitative techniques, scholars trace situations or events step-by-step to explain a social process or outcome. Recently, a methodological framework has been formulated under the label Video Data Analysis (VDA) to provide a reference point for scholars across disciplines. Our paper aims to further contribute to this effort by detailing important issues and potential challenges along the VDA research process. The paper briefly introduces VDA and the value of 21st century visuals for understanding social phenomena. It then reflects on important issues and potential challenges in five steps of conducting VDA, and formulate guidelines on how to conduct a VDA: From setting up the research, to choosing data sources, assessing their validity, to analyzing the data and presenting the findings. These reflections aim to further methodological foundations for studying situational dynamics with 21st century video data.

Keywords: videos; visual data; data collection; situational analysis; data analysis; behavior; CCTV; YouTube

1. Introduction

Since the early 2000s, researchers have developed new ways of studying social phenomena by analyzing behaviors, emotions, and interactions in real-life situations. This wave of innovative social science research is fueled by an ever-expanding pool of visual data (i.e., moving or still images). People use camcorders, mobile phone cameras, and even drones to film and photograph social life, and many public spaces are under video surveillance. The simultaneous advent of user-generated-content websites made many of these data easily accessible. In 2013, 31 percent of online adults posted a video to a website, and on YouTube alone, more than three hundred hours of video material were uploaded every minute (Anderson 2015). Many of these videos document real-life social situations and interactions, some of which are very difficult to study with established methods (e.g., rare events, such as robberies, assaults, or mass panics). Effectively, these developments crowd-source data collection. Sociologist Randall Collins (2016a) foresees a “golden age of visual sociology”.

Already, such visual data from non-laboratory settings are being increasingly employed to study phenomena in sociology, social psychology, political sciences, and criminology, among many other fields. As the common denominator, these studies use visuals as data containers

that capture human behavior, interactions, emotions, or relationships in unprecedented detail. Videos capture the sequential nature of interactions and, through their outstandingly detailed information, allow tracing the dynamics of such sequences in unprecedented detail. They enable meticulous analyses of how situational dynamics impact social outcomes and contribute substantially to our understanding of social phenomena (e.g., [Bramsen 2018](#); [Collins 2008](#); [Klusemann 2009](#); [Levine et al. 2011](#); [Nassauer 2018a](#)).

Despite their commonalities, studies lack a coherent framework that connects approaches from different fields and provides a methodological umbrella. Many authors present their analytic approach without reference to or drawing on a broader research strand, a common set of analytic tools, or quality criteria. For example, one of the most prominent examples of a sociological study of actions and interactions, [Collins \(2008\)](#) micro-sociological study on violence, uses videos and pictures of violence-threatening situations and violent outbreaks. Collins' study is a modern classic that challenges long-held explanations of the emergence of violence and human nature, but it draws criticism for being methodologically vague ([Laitin 2008](#), p. 52).

[Nassauer and Legewie \(2018\)](#) formulated a methodological framework called "Video Data Analysis" (VDA) as a first step to provide a reference point for current and future users, reviewers, and interested readers. The framework summarizes commonalities of VDA studies, situates the approach among more established methods, describes advantages and potential pitfalls of using visual data, outlines a toolkit of analytic dimensions and procedures, introduces criteria of validity, and discusses challenges and limitations. In the present paper, we build on this effort by detailing important issues and potential challenges along the research process. The paper is aimed at practitioners and scholars in relevant research fields who wish to familiarize themselves with the use of novel types of visual data to study situational dynamics.

In the first section, we reflect on how visual data can add to the analysis of situations and briefly outline VDA's key characteristics. Subsequent sections discuss important issues and potential challenges along five key steps of the research process: (1) formulating a VDA-compatible research question and setting up the research, (2) choosing data sources, (3) assessing data validity, (4) analyzing data, and (5) presenting findings. The appendix includes an overview table, as well as a list of guiding questions that help reflect on key issues during the different research steps.

2. Using Visuals to Analyze Situations

2.1. Situational Dynamics and Video Data

Since at least the early 20th century, situational dynamics have been a core interest in sociological research, as well as some areas of psychology and criminology (e.g., [Garfinkel 2005](#); [Goffman 1963](#); [Levenson and Gottman 1985](#); [Mauss 1923](#); [Reiss 1971](#)). 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?

Until the turn of the millennium, studies interested in real-life situational dynamics mostly employed participant observation of social encounters and retrospective interviews of participants to describe and explain a range of social phenomena, making key contributions to their research fields (e.g., [Humphreys and Rainwater 1975](#); [Katz 2001](#); for criminology, see ([Mastrofski et al. 2010](#))); however, there have been a number of exceptions using video, such as ([Reiss 1971](#); [Mehan 1979](#))). Since the early 2000s, an ever-increasing amount of visual data is employed to add to both of these established approaches, allowing to observe situational dynamics more reliably and in more detail (e.g., [Collins 2008](#); [Klusemann 2009](#)). These types of data and analyses are not useful when interested in phenomena including individual meaning-making, or long-term processes, such as life courses or life cycles. However, when interested in real-life situational dynamics, the unfolding of processes and human interaction and emotion, they offer specific advantages.

2.2. Advantages of Visual Data

Visual data offer three main advantages that make them a valuable addition to more established methodological approaches for analyzing situational dynamics. These advantages revolve around research opportunities, detail of analysis, and sharing of primary data.

First, visual data offer a quantum leap in access to data on real-life situational dynamics and human interaction. Second, visual data provide extremely detailed and reliable information on situational dynamics as they happened and avoid possible observer bias, post-event omission, and lack of memory, all of which may taint situational detail in participant observation and interviews (Bernard et al. 1984, p. 509f; Kissmann 2009; Knoblauch et al. 2006; Lipinski and Nelson 1974, p. 347; Vrij et al. 2014). Third, visual data allow extremely high transparency as they enable researchers to share primary data on real-life human interaction with colleagues and readers. In recent years, transparency is becoming an increasingly central criterion of good practice in social science research (e.g., ASA 2008, para. 13.05a).

2.3. Visual Data Analysis

Video Data Analysis taps into the advantages of novel video data to study situational dynamics. Nassauer and Legewie (2018) distilled the approach from related methodological fields (mainly visual studies, ethnography, psychological laboratory experiments, and multimodal interaction analysis), as well as existing studies that focus on situational dynamics with video data. Such studies have examined atrocities (Klusemann 2009), protest violence (Bramsen 2018; Nassauer 2012), drug selling (Sytsma and Piza 2018), or street fights (Levine et al. 2011; Philpot and Levine 2016; Liebst et al. 2018), emergency call centers (Fele 2008) robberies (Mosselman et al. 2018; Nassauer 2018b), and learning (Derry et al. 2010; Mehan 1979), among many other phenomena. While such VDA-related applications rely on very different theoretical underpinnings and may employ different types of quantitative or qualitative analysis, all share a focus on situational dynamics.

The most prominent among recent applications in sociology is Collins (2008) analysis of pictures and videos to study emotional dynamics in a variety of violent and near-violent situations. Collins examines the minutes and seconds before and during violent actions, especially focusing on emotion expressions identifiable in actors' facial muscles and body postures. This approach allows Collins to challenge core assumptions of conventional theories of violence by showing that situational emotions, instead of actors' prior strategies or motivations, trigger violent behaviors. Visual data are instrumental to Collins's analysis, enabling him to develop his argument and corroborate his findings. His study suggests that "there is crucial causality at the micro-level" (Collins 2016b), which can be uncovered using visual data.

A detailed discussion of VDA regarding its methodological forebears, basic principles of the approach, and example studies goes beyond the scope of this paper and has been published elsewhere (Nassauer and Legewie 2018). However, existing publications so far do not discuss key issues and challenges when conducting VDA or provide guidelines on how to conduct a VDA study. The following sections will, therefore, reflect on important issues and potential challenges when using 21st century video data for analyses in sociology, criminology, psychology, education, and beyond.

3. Discussion: Key Issues and Challenges in Video Data Analysis

In the following sections, we will discuss issues and challenges of video data analysis and formulate first guidelines on how to conduct a VDA. Many methodological basics and general good practices in social sciences apply to VDA. As other approaches, VDA requires succinct research questions, sensible case selection, a transparent and accessible documentation of data, and methods to produce meaningful and reliable findings. Discussions of these general issues fill entire handbooks (e.g., Box-Steffensmeier et al. 2008; George and Bennett 2005; King et al. 1994). For the sake of brevity, we will focus on aspects that are particular to the VDA research process. Our paper covers five key

steps of research, from setting up the VDA research, to collecting data and evaluating validity, to analyzing data and presenting findings (see Figure 1 for an overview). The order of steps 1–4 may also be switched in VDA studies, e.g., when using more inductive approaches such as Grounded Theory.

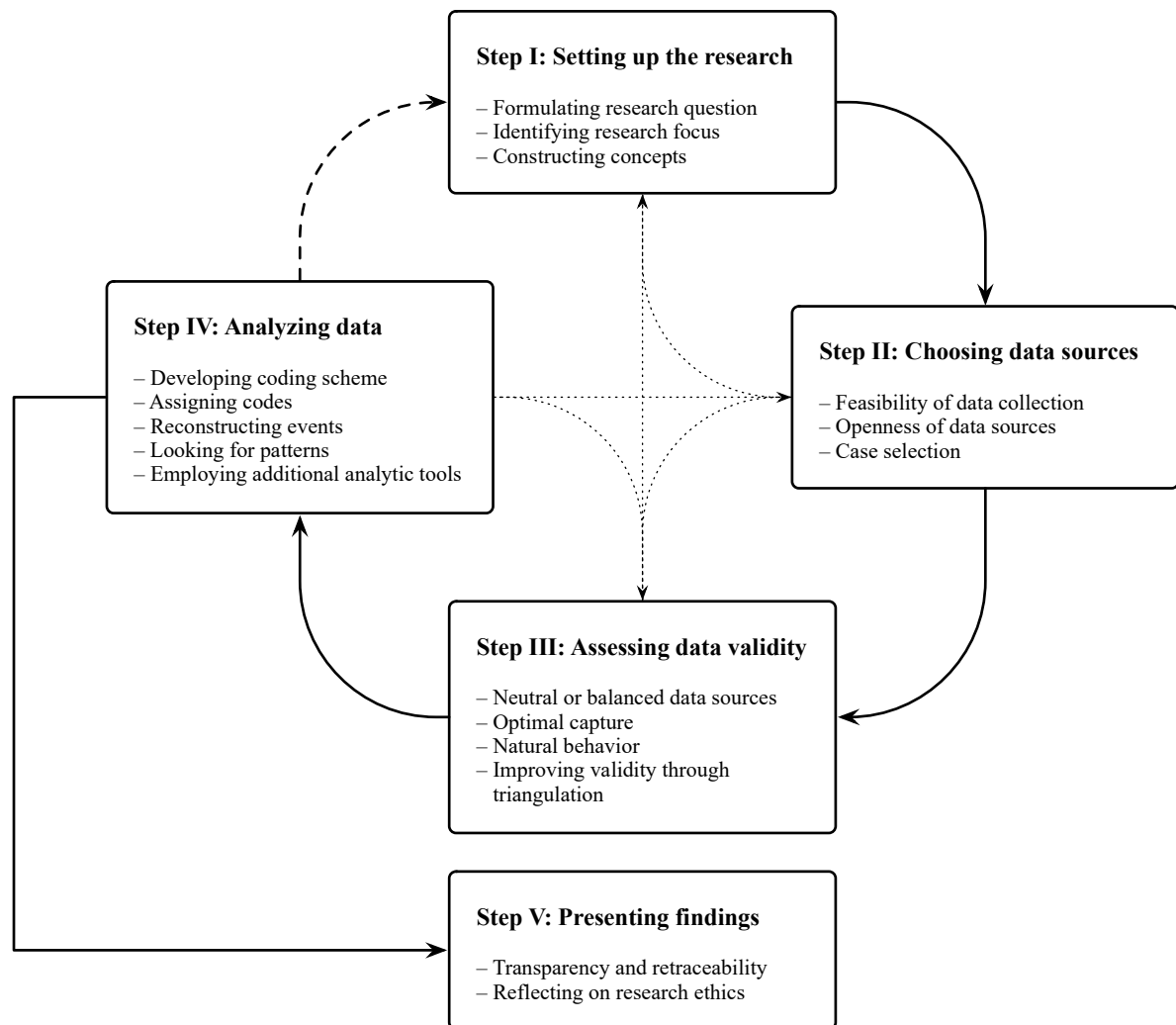


Figure 1. The Video Data Analysis (VDA) research process.

3.1. Step 1: Setting Up the Research

Key issues and challenges for a manageable and systematic VDA often arise from the research's basic setup: a clear research question, focus on key analytic dimensions, and transparent concepts.

3.1.1. Formulating a Video Data Analysis Research Question

By design, VDA is best suited for research questions focusing on interactions, emotions, or other aspects of situational dynamics. Hence, for a researcher to employ VDA, such dynamics should be expected to play an important role in explaining the phenomenon under investigation. If in doubt, one way to examine whether such a focus makes sense is for researchers to ask: How would one tackle the research topic from the perspective of researchers such as [Goffman \(1963\)](#), [Garfinkel \(2005\)](#), [Blumer \(1986\)](#), [Collins \(2005\)](#), [Katz \(2001\)](#), [Luckenbill \(1981\)](#), [Felson \(1984\)](#), [Abbott \(Abbott and Forrest 1986\)](#), or [Reicher \(2001\)](#)?

For instance, in a VDA study on bystander behavior and physical violence, [Levine et al. \(2011\)](#) ask: Do escalatory and de-escalatory actions of bystanders matter for leading from aggression to physical fights? The research question focuses on interactions between a clearly defined set of actors in

a specific type of situation. Key aspects of interest (escalatory and de-escalatory behavior of bystanders and perpetrators) are observable in the used visual data (CCTV recordings of aggressive instances). The question is thus ideally suited for a VDA. Of course, VDA can also be employed in exploratory, inductive studies. For instance, [Nassauer \(2018b\)](#) explores the ways in which situational dynamics impact store robberies by watching store robberies unfold on video and then developing a research question and key concepts grounded in these data.

Further, VDA researchers should reflect on how exactly visual data (potentially triangulated with other data; see below) will contribute to answering the research question. For instance, [Collins \(2008\)](#) finds that visual data for the first time put researchers into the situation of following first-hand emotional dynamics during violence-threatening and violence-erupting situations. They helped answer his research question on the role of situational dynamics for violent outcomes by allowing to study second by second where actors stand, how some violence-threatening situations end in violence and others did not, and which emotions actors display before violence.

3.1.2. Identifying the Research Focus

Identifying the research focus means choosing what analytic dimensions among the content of visual data should be at the center of the study. The same video can be analyzed in many different ways, focusing on different analytic dimensions ([Derry et al. 2010](#), p. 10). [Nassauer and Legewie \(2018\)](#) distinguish three analytic dimensions: facial expressions and body posture, interactions, and context. Behind each dimension lies a microcosm of information. Facial expression and body posture are any non-verbal information that a person's face and body convey. Examples include emotion expressions, body position vis-à-vis co-present actors, or calming behavior, such as actors touching their chest. Interactions refer to anything people do or say that is geared toward or affects their environment or people therein. Examples include movement in space, using items, engaging others and/or objects physically, but also gestures and verbal communication. Context means information on the physical and social setting of the situation. Examples of physical context include lighting and spaciousness. Examples of social context include the number, age, and gender of actors (see [Nassauer and Legewie \(2018\)](#) for more detail on these dimensions). Focusing on relevant dimensions of situational dynamics and constructing transparent concepts at some point during the research process is especially important in VDA because of the immense detail of information contained in visual data ([Derry et al. 2010](#), p. 6; [Norris 2004](#), p. 12). For instance, while [Collins \(2008\)](#) focuses on emotion expression (e.g., fear, anger), [Levine et al. \(2011\)](#) focus exclusively on interactions (e.g., hitting, kicking).

Due to the great detail of video data, there are many more possible foci. Researchers may, therefore, choose to reserve some time for finding a research focus, either on their own or by watching the video with a research group. If in doubt about the (preliminary) research focus, it may be helpful for researchers to use their gut response when viewing video data on the situations under investigation. Just as in ethnography, this gut response to situations can serve as a guide.

3.1.3. Constructing Concepts

VDA concepts can be developed through an inductive or deductive approach ([Derry et al. 2010](#), p. 9), applied to quantitative and qualitative analyses, and may capture broad or fine-grained actions. For instance, [Levine et al. \(2011\)](#) build concepts of escalatory or de-escalatory actions grounded in the data. They conceptualized lower-level actions, such as "escalatory behavior," inductively, breaking them down into indicator-level dimensions, such as "slapping" or "kicking," for a quantitative analysis of street fights. A VDA researcher may also apply concepts deductively, such as [Klusemann \(2009\)](#) in his qualitative analysis of the Srebrenica massacre, in which he relied on Collins' and Ekman's concepts of emotion expressions (happiness, anger, sadness, surprise, disgust, contempt) for a qualitative analysis of military negotiations. Scholars may also choose theoretical abduction as a

middle ground (Swedberg 2011, 2014), in which question, category, concept, and theory are created from data and existing theory in an iterative process.

As in other methodological approaches, concepts—whether deduced from theory or generated inductively—used in VDA should eventually entail a transparent structure with explicit connections between concept elements. Transparent structure means that clear indicators should connect abstract concepts and observations in visual data (Becker 1986, pp. 261–64). Explicit connections means that the links between abstract concepts and specific concept elements should be explicitly stated (Goertz 2006, p. 27ff). For instance, is a given specific observation (e.g., protesters throwing stones at police officers) necessary to be observed for an abstract concept (e.g., violence in demonstrations) to be seen as present? Or are there other specific observations (e.g., police using clubs to hit protesters) that would be alternative indicators of the presence of the same abstract concept? Transparent and well-structured concepts have the advantage that researchers can detect and code them reliably in the data and other scholars with access to the same data can evaluate the coding procedure or apply the same concepts to code their video data (George and Bennett 2005, p. 70; King et al. 1994, p. 2). Hence, as the goal of concept development (whether at the outset or through data analysis), we suggest that abstract notions should be connected to more palpable secondary-level dimensions, which in turn can be connected to one or more indicator-level dimensions. Goertz (2006) and many others provide useful templates for this task.

3.2. Step 2: Choosing Data Sources

The choice of data sources for VDA data collection is another key issue and holds specific challenges. Researchers can produce their own visual data by filming relevant social situations or they can use found visual data, i.e., visual data produced by third parties, such as courts, police, security camera providers, other researchers, or users of online video platforms. Regardless of the choice of data source, researchers will have to sample observations or select cases based on their research question and analytic approach, or derive their focus and questions from an exploratory data analysis (in the latter case choosing data source would be the first step during a VDA). Two key concerns when choosing VDA data sources are feasibility of data collection and openness of sources.

3.2.1. Feasibility

The different types of visual data sources each have advantages and drawbacks regarding feasibility. One important consideration is the time required for data access. If data are accessed through courts, police, or security camera providers, considerable time may be needed to prepare data access due to negotiations or bureaucratic processes. Further, researchers should be aware that, depending on the research question and the material, found video material might have limited analytic potential in some ways because the researcher cannot influence what exactly of the situation is captured (Derry et al. 2010, p. 8).

With self-recorded video material, it may be possible to produce a representative corpus of situations by checking with participants whether recordings depict a “typical” day or situation in each setting (Derry et al. 2010, p. 13). Yet, there may be negotiations involved here, too, depending on whether it is necessary to obtain permission to film. In public spaces, filming is usually allowed for research purposes if the footage captures the space itself and people move through it. If researchers film specific persons or record private spaces, they need permission from the owner(s) as well as informed consent of the people filmed (Knoblauch et al. 2006, p. 16; Wiles et al. 2011, p. 688). Also, researchers should allow ample time for planning and testing ways to record data to ensure the highest possible validity.

Depending on the research question, it may be an attractive alternative to use visual data accessible through online platforms. With the pool of visual data available online increasing by the minute, it becomes ever easier to find visual data relevant to a given research project. These data are only a mouse-click away, and thus access is often very time-efficient. This efficiency will further increase

with researchers' access to web crawling techniques. A further advantage can be that the researcher him- or herself is not present in the filmed situation, thus not influencing it as an outsider. Usually, videos uploaded online are caught by CCTV cameras or by people that are common participants in the filmed situation (e.g., friends, coworkers, bystanders). On the downside, scholars can neither influence what is captured in such data, nor which captured instances are uploaded. Moreover, specific ethical issues, such as lack of informed consent and anonymization, may apply (see step 5, see also [Legewie and Nassauer 2018](#)).

A further aspect of feasibility is the ubiquity and predictability of the situation or event under study. If the event or situation is rare, it may not be possible to rely on self-recorded data because it would take too long to collect it, and it may be impossible to select cases systematically for comparative leverage. For example, store robberies or natural disasters are too rare and unforeseeable to allow researchers to self-record them systematically. In such cases, it will usually be more feasible to rely on found visual data captured by CCTV providers or bystanders and uploaded on online video platforms.

3.2.2. Openness of Sources

For reasons of transparency and reproducibility, it is preferable to make the primary data of a study available to reviewers and readers ([Salganik 2017](#), p. 300). Visual data, in contrast to, e.g., participant observation, in theory, offers the unique opportunity to meet this high standard when analyzing social situations ([Nassauer and Legewie 2018](#)). Thus, whether opening data to the public will be possible and ethically acceptable should factor into the decision as to what types of data sources to employ in VDA.

3.2.3. Case Selection

Beyond selection of data sources, researchers select specific cases to analyze. This is an eminently important issue, and a broad literature is available that provides insightful guidance (e.g., [Gerring 2008](#); [Small 2009](#)). We will therefore limit our discussion to two brief remarks. If more potential cases are available to study a specific phenomenon than can be analyzed, a researcher may only select those cases for which data comes from reliable sources, ethical standards are not violated ([Legewie and Nassauer 2018](#)), and criteria of validity are met (see below). If these criteria do not narrow down available cases, scope conditions could be applied (focus on a specific place or timeframe), cases could be sampled randomly, or cases could be selected based on theoretical sampling. If fewer cases are available than would be useful for the analysis, researchers can try to combine types of sources (e.g., videos from Internet sources and self-recorded data).

3.3. Step 3: Assessing Data Validity

A crucial step in any research approach is to assess data validity. Validity refers to whether concepts and indicators measure what they are intended to measure, and whether the data can sustain the conclusions presented by the researcher. To assess validity in VDA, we suggest using neutral or balanced data sources, optimal capture, and natural behavior as criteria. In addition, we suggest three ways of data triangulation that can help increase validity.

While no single criterion should be considered fatal if not met completely, it is crucial to apply the criteria during each step in the research process, look for solutions when challenges arise, and, if no alternatives work, report each transgression or exclude inadequate sources or cases from the analysis ([Jordan and Henderson 1995](#), p. 56).

3.3.1. Neutral or Balanced Data Sources

Neutrality and balance of data sources may be challenged by selection bias; unintended but systematic ways in which the selected cases and/or data sources may influence findings. VDA data and data sources may be biased in a number of ways. Self-recorded visuals may suffer from reactivity, which means participants in the situation adapt their behavior because of the presence of researchers

and/or recording devices (see below for more on the issue of reactivity). Visual data by institutions, such as the police and courts, might only show events that were prosecuted or may omit certain scenes during a situation because of institutional interests. Visuals uploaded online may show more extreme cases than are the norm and researchers, therefore, need to reflect on why cases might have been uploaded.¹ To assess neutrality or balance of data sources, researchers should reflect on what biases each *source* of data may entail, how they could affect the analysis, and whether there are ways to avoid or mitigate bias. As we will discuss in detail below, triangulation of types of data sources is one way to hedge such selection bias.

3.3.2. Optimal Capture

Optimal capture describes the degree to which the data material provides information on all aspects of a situation that are relevant to a given research question (Nassauer and Legewie 2018). 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). The criterion can serve as a guideline when selecting cases and data sources, and when assessing validity of collected data. Optimal capture should be evaluated with the situation under study as the focal point, which may be captured by a single or multiple data pieces. Relevant aspects include duration, resolution, zoom, and camera angle. At what point the resolution is too low, the zoom level is inadequate, sound is relevant but missing, or the camera angle is problematic to achieve optimal capture largely depends on the analytic focus and the requirements for data it entails. For instance, analyzing facial expressions requires fairly high resolution and continuous footage of all relevant actors' faces.

3.3.3. Natural Behavior

By natural behavior we mean that participants in a situation behave the way they normally do in the type of situation under investigation. Natural behavior should not be confused with how common a behavior is, how frequently a situation occurs, or whether the data and the findings drawn from their analysis are representative. Rather, natural behavior refers to whether the fact that someone or something was recording a situation led to people acting differently than they normally do in this kind of situation.

The biggest challenge to natural behavior is reactivity. The possibility that actors adapt their behavior due to the presence of a researcher or recording device (LeCompte and Goetz 2007, p. 20ff). Nassauer and Legewie (2018) suggests assessing natural behavior with a series of three questions: (1) did participants realize they were being filmed, if so, (2) did they adapt their behavior, if so (3) how far would that pose a substantial challenge to the recorded behavior being natural (Nassauer and Legewie 2018, pp. 23–25).

(1) If people in the visual footage did not know they were recorded, reactivity can be ruled out (Pauwels 2010, p. 563). (2) If people did realize they were being filmed, there are a number of reasons that would still suggest behavior is natural: Some behavior, such as facial expressions, happens subconsciously and is thus very hard to control (Ekman et al. 1972), reactivity is low if recordings are commonplace or the situation is very important to participants (Becker 1986, p. 255f), and people tend to forget rather quickly that a camera is filming, especially if no camera operator is present (Jordan and Henderson 1995, p. 55ff; Smith et al. 1975; Christensen and Hazzard 1983). (3) If participants clearly react to cameras or other recording devices, the third question is whether such reactivity is an integral part of the situation. For instance, as protesters usually count on being filmed

¹ Not least because they may foster sample bias in online video studies, uploading practices are an interesting aspect for future research.

(Tilly and Tarrow 2006), reactivity is an essential part of most protests and can constitute natural behavior in such situations (see Nassauer 2018a).

3.3.4. Improving Validity through Triangulation

Data triangulation is a key tool to increase analytic potential and validity of VDA studies (Collins 2016a, p. 82). First, triangulation of types of data sources may help avoid selection bias and gathering as much and as diverse data as possible. Beyond the general advantage of using ample and diverse data, triangulation of data sources becomes highly important if data sources are partisan or imbalanced, for instance, because an institution or interest group with a specific agenda or stake provided visuals of the situation or event under study and may have refrained from filming certain transgressions by their own group (e.g., protesters and police, see Nassauer 2018a).

Second, triangulation of individual data pieces can allow cross-validating information between data pieces and may address incomplete captures by complementing missing information that may arise due to selective angles, actors moving out of the frame, or missing parts of the interactions. For instance, researchers may triangulate data from different CCTV cameras, different mobile phones footage of the same event, or use multiple cameras or different zooms when self-recording.

Third, triangulation of videos with other types of data, such as interviews, documents, or field notes from participant observation, can be especially helpful for gathering information on context or actors' interpretations of the situation, for instance on actors' interpretation of the event, their knowledge of what is happening, or the spread of rumors (e.g., protesters and police, see Nassauer 2018a). Scholars may also interview participants of an event (whether protest event, school event, or other), ask them for their videos, and review the material with them to gain further insights that can inform their VDA (e.g., Bramsen 2018). All three suggested strategies provide opportunities for VDA researchers to increase validity, depending on their respective research project, analytic focus, and research question.

3.4. Step 4: Analyzing Data

Often, the main goal of VDA is to determine whether there is a connection between situational factors and the outcome of interest. In some cases, this means to test a hypothesis or theory formulated beforehand; e.g., whether situational patterns are necessary for violence emergence (Bramsen 2018; Klusemann 2009). In other cases, the approach will be more inductive, analyzing data to explore whether situational dynamics may explain when and how a given outcome occurs, e.g., how far situational pathways are relevant to a successful crime (Nassauer 2018b).

How to analyze data in VDA depends on issues, such as the research question and the theoretical frame. Nonetheless, there are some general issues of data analysis that will often matter in VDA, which we will discuss in the following section. Figure 2 provides an overview of a possible VDA workflow and the substeps involved: developing a coding scheme, assigning codes, reconstructing events, looking for patterns, potentially using additional analytic tools, and presenting findings.

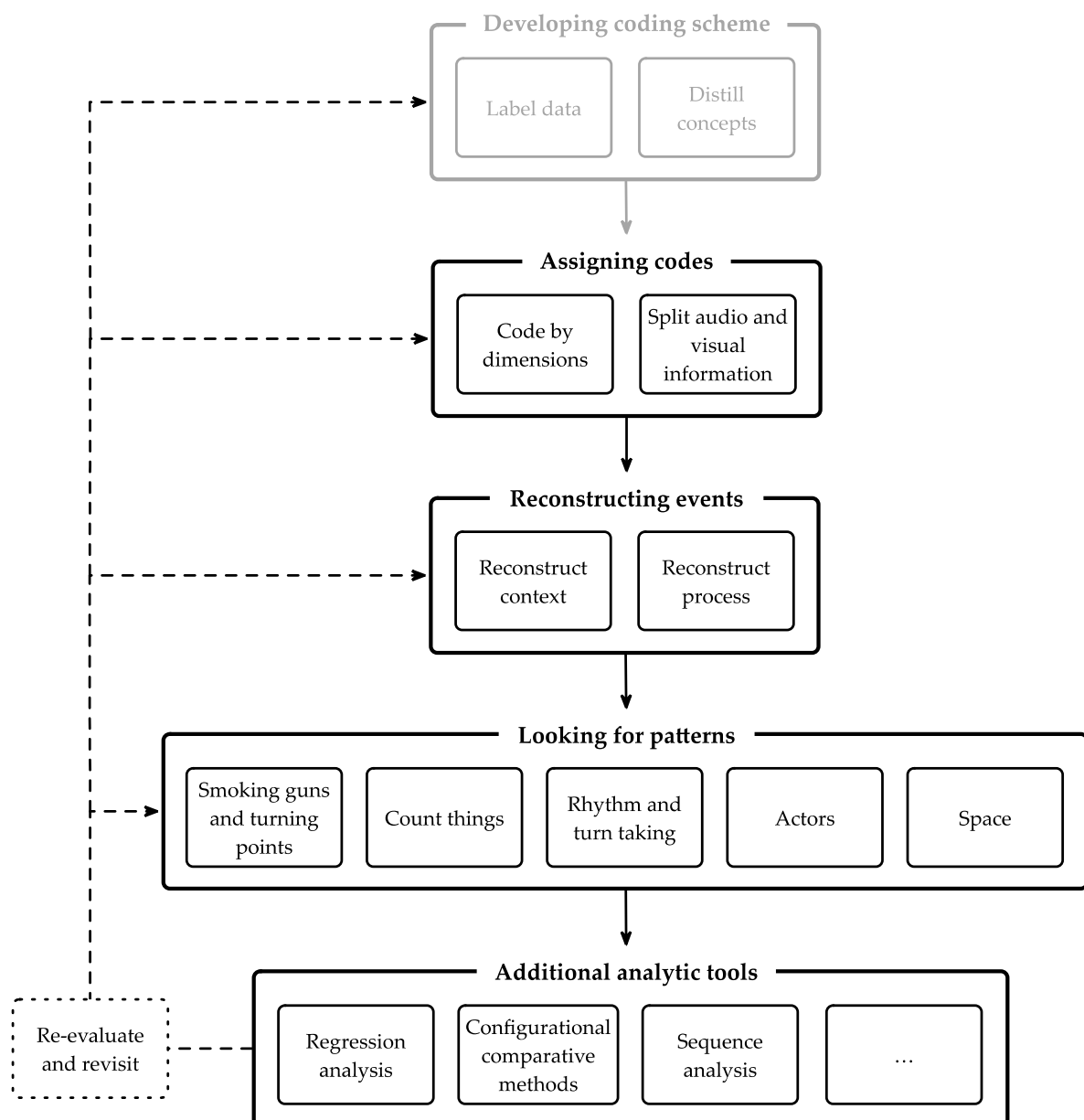


Figure 2. Overview of data analysis in VDA.

3.4.1. Developing a Coding Scheme: Label Data and Distill Concepts

Like other analytic approaches, VDA works by using a set of concepts to make sense of a data set with the goal to answer a certain research question. Some researchers will follow a deductive approach, in which concepts will be formulated based on existing theories and empirical studies. In this scenario, developing concepts and a coding scheme is not a part of data analysis. Researchers using a deductive approach can start with assigning codes to the data right away.

If there is little existing knowledge on the topic of research or researchers wish to retain some flexibility in how to formulate their concepts and coding scheme, they may choose an inductive approach, making developing concepts and a coding scheme part of the analysis. There are excellent guidebooks for how to develop concepts from the data (e.g., [Strauss and Corbin 1998](#)), so some very brief remarks will suffice here.

Roughly, the process of developing concepts and a coding scheme can be split into a first phase of intensive, detailed labeling of information in a data piece, and a second phase of grouping labels

together into categories and dimensions that can be connected to form concepts. By going back and forth between the two phases in an iterative process, researchers eventually construct concepts and a coding scheme from the data. They should be made explicit using concept structures as discussed above.

In the first phase, 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" (Levine et al. 2011, p. 408). 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.

3.4.2. Assigning Codes: Code by Dimensions and Split Audio and Visual Information

Codes, whether they are used deductively or generated inductively, synthesize a section of a data piece and thus function as landmarks helping researchers to retrieve information from their data. Assigning codes to data means to assign one or several codes to a section of a data piece, thereby indicating the relevant content this section contains. In principle, coding visuals can be done by hand, but software packages, such as Atlas.ti, Boris, Noldus Observer, NVivo, or MAXQDA, greatly facilitate systematic coding as well as later access to the coded data material.

Due to the complexity of situational dynamics, VDA researchers may want to facilitate the coding process. One way of doing so is to slow down the playing speed of the video. Moreover, researchers may do several coding iterations, each with a different focus. One way to do this is to focus on a single dimension at a time, such as face and body postures (Norris 2004, p. 12). If several aspects of a given dimension are relevant, it may even make sense to further subdivide the coding steps (e.g., facial expressions first, body posture later).

Another possibility to reduce complexity is to focus on visual and audio information separately, in case the latter is available. Beyond reducing complexity, this separation can also help ascertain whether information from audio and video channels corroborate or contradict each other regarding how a situation unfolds, and how each channel influences the outcome. Moreover, audio information may serve to cross-validate data. It may reveal the presence of other people or external influences (e.g., traffic noise or music) not captured on film.

The process of assigning codes to the data essentially breaks down situations to its core aspects as relevant to a given researcher or research project.² From here, researchers can start an analytic reconstruction of events, which in turn paves the way for a systematic search for patterns (or lack thereof) in situational dynamics that may (or may not) explain the outcome under study.

² There are excellent guidebooks for how to develop concepts and coding schemes from the data, and how to use them to code data (e.g., Strauss and Corbin 1998). The methodological principles when using video data do not differ substantially from other data formats.

3.4.3. Reconstructing Events

Based on the coded data, researchers can reconstruct in detail the situations or events under study by focusing on who did what, when, and where. This reconstruction should include context as well as process and can be documented in the form of a narrative account of the situation or event.

First, the context of a situation can play an important role in how that situation unfolds. Context includes physical properties, social aspects, and broader social or political events. Such information can be crucial for interpreting situational dynamics. For instance, at a memorial service for the January 2015 Paris terrorist attack in November of the same year, attendees reacted in panic upon hearing the sound of what turned out to be firecrackers going off in the vicinity ([Channel 4 News 2015](#)). Missing context information would render this situation unintelligible by researchers.

Information on relevant physical characteristics of space may include the expanse, topography, lighting, and weather conditions. Information on relevant objects may include furniture, items actors use during the situation, and devices, such as television screens or loudspeakers, among many other things. Information on actors may include the overall number of actors present, information on relations between actors, and demographic characteristics such as age, gender, and race (if there is reliable information available on such aspects). Information on the social context may include how access to the physical space is organized, what types of behavior is usually seen as acceptable, and whether actors gathered with a specific goal or purpose.

For some of these aspects, visual data contain the required information. For others, researchers have to try obtaining additional data of other types, if such information is relevant to their research question. Possible sources for such information could be news reports, interviews with participants, or any number of other non-visual data. We suggest producing an information sheet for every situation or event studied in a research project that details relevant information on the social and physical context as well as missing information that could limit the analysis. Such information sheets will help VDA researchers during the research process, but also serve as a useful resource for readers.

Second, processes can be reconstructed using video data. In order to analyze a captured social situation to assess whether its situational syntax impacts a specific social outcome, VDA researchers usually produce some kind of step-by-step processual reconstruction for each case—what happened when, following what, and leading to which subsequent development. Reconstructing the temporal sequence of a situation in this way lays the groundwork for further analysis and is often a fundamental part of working with visual data in VDA. Analyses usually produce a matrix of time and actions or sub-events, a diagram of movements, or some other type of representation of process within a situation. For example, relevant types of actions or sub-events could occupy the columns of a matrix, and rows mark progression over time. In each cell, a researcher can indicate which action or sub-event occurred and, if so, provide further information. Such an approach was taken by [Klusemann \(2009\)](#) when reconstructing emotion expressions and speech of UN commander Karremans and Serbian general Mladic during an eight-hour video of their negotiations prior to the 1995 Srebrenica massacre (Figure 3a).

[Levine et al. \(2011\)](#) take a different approach. The authors follow the development of aggressive situations step-by-step and focus on the influence of bystander actions. Shown in Figure 3b, the authors construct a sequential action tree for each situation, showing the succession of bystanders' escalatory and conciliatory behaviors after an initial escalatory behavior by a perpetrator ([Levine et al. 2011](#), p. 410).

If the situation or event under study takes place in a large area, establishing a temporal sequence can be insufficient to produce a seamless case reconstruction. It may be necessary to locate each behavior or sub-event in space, as well. For instance, [Nassauer \(2016\)](#) studies temporally and spatially extensive protest events. Similar to the generic example in Figure 3c, the author located each data piece in space and time before being able to study the situational dynamic and identify connections

between interactions and participants’ emotional states. To achieve this, Nassauer used documented protest routes, police radio logs, and Google Maps Street View.

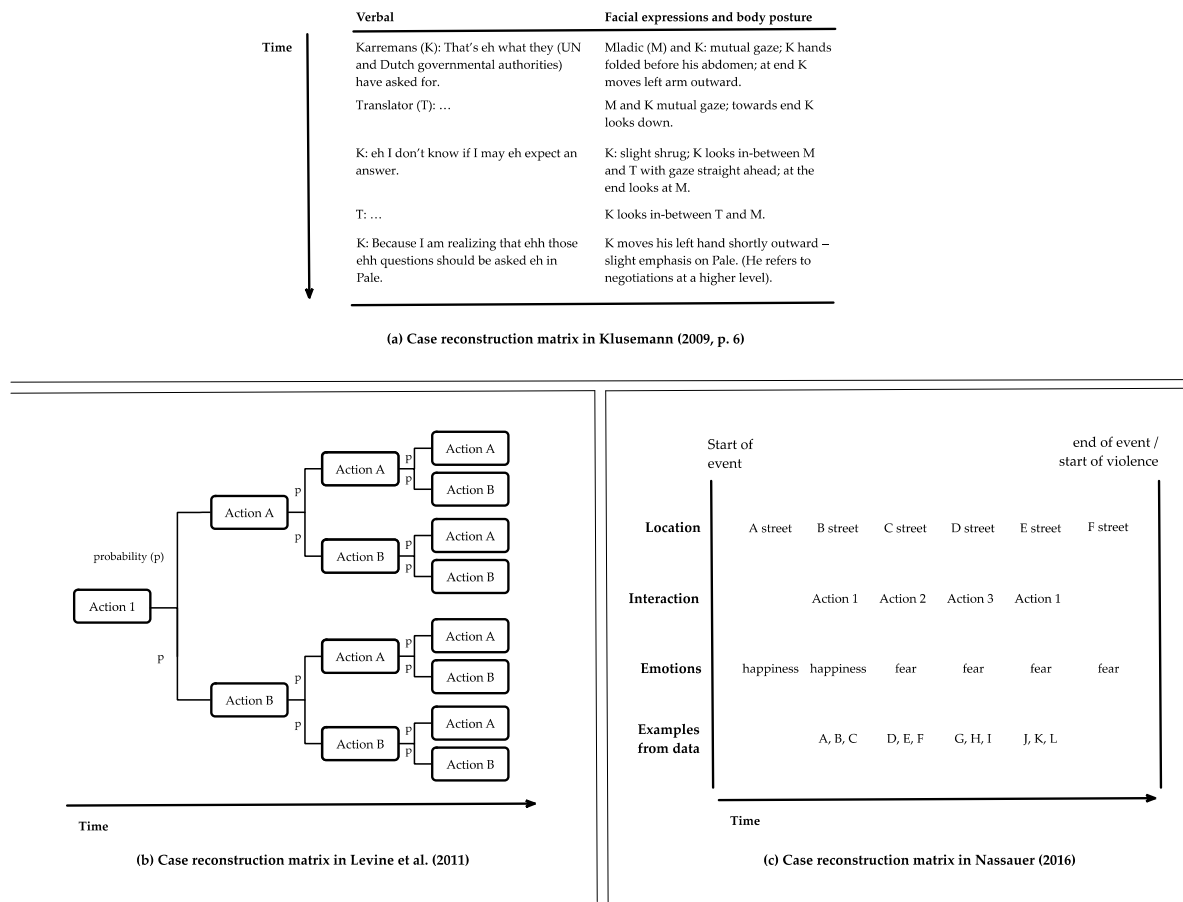


Figure 3. Case reconstruction matrices.

3.4.4. Looking for Patterns

Reconstructing context and process tells us what happened, when and who was involved. However, this chronologic order is not equivalent to the sequential order of actions, reactions, and interactions during a situation. Rather, reconstruction serves as the basis for asking: What situational path led (through one or several chains of linked actions and stages) from the start of the situation to the occurrence, or absence, of the outcome under study (see Blatter 2012, p. 5)? Tackling such questions requires looking for patterns within (and potentially also across) cases. In this sense, VDA operates similarly to causal process tracing and its understanding of processes as a temporal sequence of linked elements (Blatter 2012; Mahoney 2012), VDA, however, has a clear focus on the micro-level of situations.

The following perspectives may help to understand situational patterns and identify key factors and mechanisms: (1) smoking guns and turning points, (2) count things, (3) rhythm and turn-taking, (4) actors, and (5) space. These perspectives should be understood as a non-exclusive, non-exhaustive list. Depending on the research question and data at hand, other ways to look for patterns in situational dynamics may well prove fruitful.³ In the following, we will discuss each of these aspects in turn.

³ As a general note, it may be fruitful to examine video data for patterns by switching between normal and double video run speed. The faster run speed can help pulling back from single actions and reactions to get a view of the flow of a situation.

First, looking for smoking guns and turning points means to scrutinize the tempo-spatial sequence of actions and behaviors for links of cause and consequence. A first step in identifying such links is to look for smoking-gun observations (Blatter 2012, p. 16), temporally and spatially adjacent actions or other occurrences that are linked by a clear dynamic or mechanism.

Extending this notion beyond immediately adjacent actions or other occurrences, it can be fruitful to pay attention to changes in the analytic dimensions of interest, and whether changes in those dimensions are connected. This may help identify pivotal moments in which clear shifts in the situational dynamic occur, such as turning points, critical junctures, or windows of opportunity (Goertz and Levy 2007). Such turning points may be the sudden transformation in relevant sub-dimensions of face and body posture or interaction. Sudden changes, if they occur, are usually discernable from context and from studying the video closely. Once such moments have been identified, researchers can work forward and backward from these shifts to identify their causes and/or effects. For instance, Klusemann (2009) studies General Mladic and Commander Karremans' respective emotional states (Figure 3a) by analyzing the two protagonists' actions and reactions. The author pays special attention to the rhythmicity of the situation, how participants take turns during their interaction, and how disruptions to this rhythmicity eventually lead Mladic to develop emotional dominance over Karremanns.

Second, and in addition to identifying smoking guns and turning points, it may be interesting to look at the intensity, frequency, and duration of certain actions, behaviors, and interactions (Gross and Levenson 1993, p. 973f). Intensity refers to a rating of how intense an interaction was. For instance, to assess how intense an act of violence was, researchers could count the number of actors involved, the number of blows, the power of those actions (a slap or a strong blow), or the severity of resulting injuries.

Frequency refers to how often a behavior or interaction occurred during a situation. For instance, Alibali and Nathan (2007) count the frequency of teacher gestures during classroom interactions and find that teachers use gestures more frequently when explaining a new topic, which helps students grasp core ideas.

Duration refers to how long a behavior or interaction lasted. For instance, Collins (2008) shows that violent altercations tend to be very short, which supports his claim that humans generally avoid violence.

Third, VDA researchers can further scrutinize if and how participants follow a rhythm or take turns during interactions. They can do so by assessing the length of each interaction, or the pauses between them. Doing so can provide information on how behaviors and interactions in a situation are organized and how participants react if that organization is disrupted (Jordan and Henderson 1995, p. 59ff). For instance, Collins (2005, 2008) studies how the human tendency to automatically mimic and synchronize movements with each other, and thereby converge emotionally, makes it more difficult to use violence.

A fourth area to scrutinize is key actors. First, the analysis of specific actors can be done in isolation, i.e., as if the actor were alone in the situation. For instance, if a researcher analyses bar fights, it makes sense to pay specific attention to the actor who throws the first punch, following that person closely in her/his journey through the situation. In a second step, the actors' behaviors can be studied in connection to other actors in physical and temporal proximity. In a third step, the impact of the actors' behavior on the overall situational dynamic can be assessed, e.g., how far actions, or emotion expressions were actions and/or reactions to another actor's behaviors or expressions. For instance, Nassauer (2018b) (see also guiding questions on this research step, Appendix A) coded the actions by each robber and clerk for 20 robberies in Atlas.ti. As one step of analysis, she then checked whether one actor regularly was first to initiate a new action to which the other responded, and whether patterns of action-reaction between the two impacted success or failure of the robbery. Another option is to focus specifically on actors from a certain group (e.g., protesters or police) or those who have a certain social role (e.g., employees and customers) within the situation under investigation.

It may prove useful to pay explicit attention to inanimate objects and whether actors use or relate to them during a situation. Inanimate objects, such as tools and devices, TV screens, or audio speakers, can play a role in how a situation develops because they often mediate actions and interactions between people. In [Nassauer \(2018b\)](#) analysis of store robberies caught on CCTV, store items are often mediating interactions and actors' roles.

Lastly, space can be a focus during video data analysis. The physical properties of a space in which a situation takes place can impact how that situation unfolds. For instance, in her VDA of protest violence in Tunisia and Bahrain, [Bramsen \(2018\)](#) shows how spatial proximity or distance, as well as elevation of the body (on a chair or building) can impact situational patterns of violence. It may also be of interest if a situation involves changes in the immediate space they unfold in, for instance, because participants move in the course of the situation. As the physical properties of the space change, so may situational dynamics and rhythms. Further, researchers may focus on spatial layers of the situation by looking at behavior, actions, and interactions in the foreground and the background of a situation. Lastly, space itself may have a specific meaning during situations. For instance, if a customer enters the space behind the counter in a shop or bar, this will likely be interpreted as a transgression by employees or owners, and may well change the situational dynamic.

3.4.5. Additional Analytic Tools

In principle, VDA's detailed case reconstructions and analyses of patterns may be combined with any method of data analysis that helps identifying patterns in a formalized and systematic way. Regression analysis ([Levine et al. 2011](#)), configurational comparative methods (e.g., qualitative comparative analysis; ([Ragin 1987](#)); see [Legewie \(2013\)](#) for a brief introduction), and sequence analysis ([Abbott and Forrest 1986](#); [Aisenbrey and Fasang 2010](#)) can be especially fruitful additions to VDA-type analyses.

3.5. Step 5: Presentation of Video Data Analysis Findings

3.5.1. Transparency and Retracement

When presenting findings, transparency is a key criterion. While optimal capture and natural behavior ensure validity of the data, transparent presentation of the findings is vital for retraceability in VDA. Retracement refers to the extent to which the analytic steps in an analytic procedure are made explicit and can be evaluated and possibly reproduced by reviewers and readers.

To some extent, transparency depends on the type of data source used for the analysis. Confidential data attained through courts, police, or security camera providers, just like participant observation and most interview data, can only be analyzed by the researcher her- or himself. They limit transparency, as data cannot easily be shared or made public. When using visual data uploaded online, a researcher can freely share the data analyzed, show videos at conferences, or provide links to the data in a research article. She or he may even use screen-shots in a publication or present the video at conferences. For instance, [Nassauer \(2018b\)](#) analyzes robberies recorded on CCTV and uploaded on YouTube and provides links to the videos used as primary data, which can be examined by all readers with an Internet access. Hence, every reader interested in the analysis is able to retrace and evaluate it. This creates a level of transparency that, to our knowledge, is difficult to achieve in any other type of data in social science research when it comes to analyzing situational dynamics.

3.5.2. Reflecting Research Ethics

At the same time, researchers should consider legal and ethical reasons when publishing VDA findings and making data accessible to colleagues and readers ([Legewie and Nassauer 2018](#)). Self-recorded visual data will usually require asking for consent from the people filmed. Consent to data being published is not automatically covered by the consent to being filmed. Data obtained

through courts, police, or security camera providers may later be subject to more strict privacy restrictions, which may inhibit sharing the material with other researchers or the public.

With data accessed online, researchers may be able to share them freely but have to be mindful of intellectual property laws and privacy rights of the individuals depicted in the data. Regarding the latter, researchers should consider the type of context from which they retrieved the data, the type of context and behavior the data depicts, and whether there may be potential harm to people depicted in the videos (Legewie and Nassauer 2018). Moreover, it may remain unclear whether the persons filmed in the videos uploaded indeed agreed to be filmed. Making the recordings available, e.g., in a data repository, on conferences, or in an exhibition requires additional consent (for a detailed discussion of ethical issues in online video research, see (Wiles et al. 2011)) and needs to be reflected upon when considering such data sources for VDA.

4. Conclusions

With novel types of visuals and the opportunities they entail for understanding human action and social life, we can expect situational research in various disciplines, from sociology, criminology, psychology, education research and other related fields to blossom. Hence, VDA studies are likely to increase in the near future.

This paper provided an in-depth discussion of key issues and challenges in Video Data Analysis (VDA) along with various guiding steps of the research process, from setting up the research to choosing data and assessing data validity, to analyzing data, and publishing findings. This in-depth discussion aims to facilitate valid, reliable, and transparent analysis for practitioners, reviewers and interested researchers across the social sciences.

Author Contributions: The authors contributed in equal parts to all aspects of this paper.

Funding: This research received no external funding.

Acknowledgments: The authors would like to express their gratitude to the participants of the 2015 video analysis workshop in Copenhagen (in particular Wim Bernasco, Isabel Bramsen, Marie Bruvik Heinskou, Mark Levine, Marie Lindegaard, Lasse Liebst, Richard Philpot, Poul Poder, and Don Weenink), participants at the *American Sociological Association* Annual Meeting, Seattle 2016, Section “Methodology—Advancement in Observing and Modeling Social Processes”, and at the *American Society of Criminology* Annual Meeting, Atlanta 2018, Section: “Methodology—Advances in Qualitative Methods”. Further, we would like to thank the anonymous reviewers for their thoughtful comments and suggestions. We acknowledge support by the German Research Foundation and the Open Access Publication Fund of the Freie Universität Berlin.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Overview Table of Key Issues and Challenges in five-step VDA Research Process.

Step	Topics	Guiding Questions
1. Setting up the research	Formulating a Research Question	Are situational dynamics expected to play an important role in explaining the phenomenon under study? Are key aspects of research question observable in visual data? Will visual data be able to answer my research question? Does my research question focus on actions, interactions, behavior, emotion expression, or other aspects of situational dynamics?
	Identifying research focus	Which analytic dimensions are relevant to the research? Do all interactions by all actors need to be coded, or only specific actions or specific actors? Are facial expressions relevant or body postures as well? How broad or fine-grained does the level of analysis need to be?
	Constructing concepts	Can I form or use transparent concepts (inductive, deductive, or abductive approach)? Do the dimensions capture what the concepts seek to describe? Do the respective indicators allow direct coding of data? How are the levels within a concept connected?
2. Choosing data sources	Feasibility	How time-consuming will data access be? Are permission and informed consent required to collect/use the data? How foreseeable and frequent is the situation or event?
	Openness	How important is it to make the primary data available to other researchers during the research process or to readers later on? What ethical and legal considerations might arise when sharing data?
	Case selection	Are more cases available than can be analyzed given the constraints of the research at hand? Are less cases available than needed for the research question at hand?
3. Assessing data validity	Neutral or balanced data sources	Who recorded a data piece and what was this person's or organization's role in the event? Does that role suggest an unbiased recording of the event? Is it possible to include data from different stakeholders in order to hedge against potential bias?
	Optimal capture	Does the data piece or patchwork of data capture all relevant aspects of the situation or event, or is additional contextual information needed? Did the relevant interaction begin after the recording began? Is every second of a relevant situation, every participant, all space, and every relevant channel covered? Are there any out-of-frame interactions or frame-to-out-of-frame interactions? Do the solution, zoom, and angle allow analyzing the dimensions necessary to answer the research question?
	Natural behavior	Is there information about actors' reactivity; e.g., by reflecting on the type of event, by actors' recorded behavior, by analyzing document data? Where actors able to see the camera? Was their behavior too routine or too important to be adapted to camera presence and if so, does this impact the analysis? Could single actors, a group of actors, or the entire situation or event be scripted and, if so, how does this affect the analysis?
	Improving validity through triangulating types, sources, or pieces of data	Does the data set show a lack of neutral or balance data sources? What types of data sources should a data set contain in order to avoid bias? Does the data set show sub-optimal capture? Is it possible to use multiple data pieces depicting the same situations to compensate for deficiencies of single data pieces? Does the data set show a lack of information regarding that it cannot be compensated by visual data? Could other data sources be used to combine their respective strengths? Is it possible to use data pieces from different types of data to cross-validate findings?

Table A1. Cont.

Step	Topics	Guiding Questions
4. Analyzing data	Develop a coding scheme	Inductive: Does a coding scheme already exist in the literature that seems useful for coding? Does it make sense to apply all the categories in that scheme to all actors, or just part of them? Is the coding scheme transparent and detailed? Deductive: Is it useful to develop a coding scheme from the data? Which elements need to be captured? Can I assign labels to the data and then form overarching categories for these labels? Can these categories and dimensions later be translated into concepts?
	Assign codes	Does it make sense to split coding into multiple iterations to reduce complexity for each criterion? If the audio track is available: How does it relate to the visual? Can additional information be gained by listening only to the sound without video and then only watching the video without sound? Does it make sense to use software for coding both visual and document data (e.g., Atlas.ti)? Does it make sense to use software that detects facial expressions or recognizes faces in visual data (e.g., Noldus Observer)?
	Reconstruct events	Does the case reconstruction based on the coded data provide a seamless account of each situation or event from start to finish, covering all relevant aspects? What information can I gather on the social or physical context of the event or process under study? How can I gather a processual reconstruction of the process or event under study? Can I identify a situational syntax by producing a matrix of time and actors or sub-events?
	Looking for patterns	Can I identify patterns in situational dynamics that can help explain the process or outcome? Did a situational path led (through one or several chains of causally linked actions and stages) from the start of the situation to the occurrence, or absence, of the outcome under study? Are there clear patterns of association between the explanatory factors and the outcome, and do these patterns apply to all or specific subsets of cases? Identifying smoking guns: What caused shifts in situational dynamics, and how does a given key moment impact the subsequent development of a situation? Turning points: Are there pronounced changes in relevant analytic dimensions during a situation or event? Are there pivotal moments in which the direction of the situational flow changes drastically? Space: Is a focus on space useful to identify patterns in the data? Are special means required to locate interactions in space? Are physical properties of space relevant to the situation or event? Actors: Is a focus on actors useful to identify patterns in the data? How does the situation unfold from the perspective of key actors? Additional Tools: Are additional analytic procedures useful to identify systematic links between explanatory factors and the outcome of interest?
5. Presentation of Findings	Transparency and Retraceability	Is a transparent presentation of the data and findings possible to ensure reliability? How important is transparency to a researcher? Which advantages or disadvantages do different data sources imply the ability to share the findings? How can I make analytic methods, and analytic procedures transparent so that other researchers can evaluate or possible reproduce the analysis?
	Ethical Issues	Do ethical considerations apply? Were study subjects able to give consent to a study? Was the video captured at a public or private event? e.g., does reprinting stills of the video violate privacy rights? Does posting links to analyzed raw data online pose potential harm to study subjects?

References

- Abbott, Andrew, and John Forrest. 1986. Optimal Matching Methods for Historical Sequences. *The Journal of Interdisciplinary History* 16: 471–94. [CrossRef]
- Aisenbrey, Silke, and Anette E. Fasang. 2010. New Life for Old Ideas: The ‘Second Wave’ of Sequence Analysis Bringing the ‘Course’ Back into the Life Course. *Sociological Methods & Research* 38: 420–62. [CrossRef]
- Alibali, Martha W., and Mitchell J. Nathan. 2007. Teachers’ Gestures as a Means of Scaffolding Students’ Understanding: Evidence from an Early Algebra Lesson. In *Video Research in the Learning Sciences*. Edited by Ricki Goldman, Roy Pea, Brigid Barron and Sharon J. Derry. New York and London: Routledge.
- Anderson, Monica. 2015. 5 Facts about Online Video, for YouTube’s 10th Birthday. *Pew Research Center (blog)*. February 12. Available online: <http://www.pewresearch.org/fact-tank/2015/02/12/5-facts-about-online-video-for-youtubes-10th-birthday/> (accessed on 15 December 2018).
- ASA. 2008. *Code of Ethics of the ASA Committee on Professional Ethics*. Washington, DC: American Sociological Association (ASA).
- Becker, Howard Saul. 1986. *Doing Things Together: Selected Papers*. Evanston: Northwestern University Press.
- Bernard, H. Russell, Peter Killworth, David Kronenfeld, and Lee Sailer. 1984. The Problem of Informant Accuracy: The Validity of Retrospective Data. *Annual Review of Anthropology* 13: 495–517. [CrossRef]
- Blatter, Joachim. 2012. *Ontological and Epistemological Foundations of Causal-Process Tracing: Configurational Thinking and Timing*. Antwerp: ICPR Joint Sessions.
- Blumer, Herbert. 1986. *Symbolic Interactionism: Perspective and Method*. Oakland: University of California Press.
- Box-Steffensmeier, Janet M., Henry E. Brady, and David Collier, eds. 2008. *The Oxford Handbook of Political Methodology*. Oxford: Oxford University Press.
- Bramsen, Isabel. 2018. How Violence Happens (or Not): Situational Conditions of Violence and Nonviolence in Bahrain, Tunisia and Syria. *Psychology of Violence* 8: 305. [CrossRef]
- Channel 4 News. 2015. Matt Frei’s Live Report as Crowds Run in Panic. Available online: <https://www.youtube.com/watch?v=sq9vpgaBgzY> (accessed on 10 January 2019).
- Christensen, Andrew, and Ann Hazzard. 1983. Reactive Effects during Naturalistic Observation of Families. *Behavioral Assessment* 5: 349–362.
- Collins, Randall. 2005. *Interaction Ritual Chains*. Princeton and Oxford: Princeton University Press.
- Collins, Randall. 2008. *Violence: A Micro-Sociological Theory*. Princeton: Princeton University Press.
- Collins, Randall. 2016a. On Visual Methods and the Growth of Micro-Interactional Sociology: Interview to Professor Randall Collins (with Uliano Conti). In *Lo Spazio Del Visuale. Manuale Sull’utilizzo Dell’immagine Nella Ricerca Sociale*. Edited by Uliano Conti. Rome: Armando Editore, pp. 81–94.
- Collins, Randall. 2016b. The Sociological Eye: What Has Micro-Sociology Accomplished? April 17. Available online: <http://sociological-eye.blogspot.com/2016/04/what-has-micro-sociology-accomplished.html> (accessed on 20 December 2018).
- Derry, Sharon J., Roy D. Pea, Brigid Barron, Randi A. Engle, Frederick Erickson, Ricki Goldman, Rogers Hall, Timothy Koschmann, Jay L. Lemke, Miriam Sherin, and et al. 2010. Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics. *Journal of the Learning Sciences* 19: 3–53. [CrossRef]
- Ekman, Paul, Wallace F. Friesen, and Phoebe Ellsworth. 1972. *Emotion in the Human Face*. New York: Pergamon.
- Fele, Giolo. 2008. The Collaborative Production of Responses and Dispatching on the Radio: Video Analysis in a Medical Emergency Call Center. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* 9: 1175.
- Felson, Richard B. 1984. Patterns of Aggressive Social Interaction. In *Social Psychology of Aggression: From Individual Behavior to Social Interaction*. Edited by Amélie Mummendey. Berlin and Heidelberg: Springer, pp. 107–26.
- Garfinkel, Harold. 2005. A Conception of and Experiments with “Trust” as a Condition of Concerted Stable Actions. In *The Production of Reality: Essays and Readings on Social Interaction*. Edited by Jodi A. O’Brien. Thousand Oaks: Sage Publications, pp. 370–80.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge: The MIT Press.
- Gerring, John. 2008. Case Selection for Case-Study Analysis: Qualitative and Quantitative Techniques. In *The Oxford Handbook of Political Methodology*. Edited by Janet M. Box-Steffensmeier, Henry E. Brady and David Collier. Oxford: Oxford Handbooks Online.

- Goertz, Gary. 2006. *Social Science Concepts: A User's Guide*. Princeton: Princeton University Press.
- Goertz, Gary, and Jack S. Levy. 2007. Causal Explanation, Necessary Conditions, and Case Studies. In *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals*. Edited by Jack S. Levy and Gary Goertz. New York: Routledge.
- Goffman, Erving. 1963. *Behavior in Public Places*. New York: The Free Press.
- Gross, James J., and Robert W. Levenson. 1993. Emotional Suppression: Physiology, Self-Report, and Expressive Behavior. *Journal of Personality and Social Psychology* 64: 970–86. [\[CrossRef\]](#)
- Humphreys, Laud, and Lee Rainwater. 1975. *Tearoom Trade: Impersonal Sex in Public Places*, 2nd ed. New York: Aldine Transaction.
- Jordan, Brigitte, and Austin Henderson. 1995. Interaction Analysis: Foundations and Practice. *The Journal of the Learning Sciences* 4: 39–103. [\[CrossRef\]](#)
- Katz, Jack. 2001. *How Emotions Work*, new ed. Chicago: University of Chicago Press.
- King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry*. Princeton: Princeton University Press.
- Kissmann, Ulrike Tikvah. 2009. *Video Interaction Analysis*. Frankfurt am Main: Peter Lang.
- Klusemann, Stefan. 2009. Atrocities and Confrontational Tension. *Frontiers in Behavioral Neuroscience* 3: 1–10. [\[CrossRef\]](#)
- Knoblauch, Hubert, Bernt Schnettler, and Jürgen Raab. 2006. Video Analysis. Methodological Aspects of Interpretative Audiovisual Analysis in Social Research. In *Video-Aanalysis: Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*. Edited by Hubert Knoblauch, Bernt Schnettler, Jürgen Raab and Hans-Georg Soeffner. Frankfurt am Main: Peter Lang, pp. 9–28.
- Laitin, David D. 2008. Confronting Violence Face to Face. *Science* 320: 51–52. [\[CrossRef\]](#)
- LeCompte, Margaret D., and Judith Preissle Goetz. 2007. Problems of Reliability and Validity in Ethnographic Research. In *Qualitative Research Volume Two. Quality Issues in Qualitative Research*. Edited by Alan Bryman. Sage Benchmarks in Social Research Methods. Thousand Oaks: Sage Publications, Inc., pp. 3–39.
- Legewie, Nicolas. 2013. An Introduction to Applied Data Analysis with Qualitative Comparative Analysis (QCA). *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* 14: 1961.
- 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. [\[CrossRef\]](#)
- Levenson, Robert W., and John M. Gottman. 1985. Physiological and Affective Predictors of Change in Relationship Satisfaction. *Journal of Personality and Social Psychology* 49: 85–94. [\[CrossRef\]](#) [\[PubMed\]](#)
- Levine, Mark, Paul J. Taylor, and Rachel Best. 2011. Third Parties, Violence, and Conflict Resolution: The Role of Group Size and Collective Action in the Microregulation of Violence. *Psychological Science* 22: 406–12. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liebst, Lasse Suonperä, Marie Bruvik Heinskou, and Peter Ejbye-Ernst. 2018. On the Actual Risk of Bystander Intervention: A Statistical Study Based on Naturally Occurring Violent Emergencies. *Journal of Research in Crime and Delinquency* 55: 27–50. [\[CrossRef\]](#)
- Lipinski, David, and Rosemary Nelson. 1974. Problems in the Use of Naturalistic Observation as a Means of Behavioral Assessment. *Behavior Therapy* 5: 341–51. [\[CrossRef\]](#)
- Luckenbill, David F. 1981. Generating Compliance: The Case of Robbery. *Urban Life* 10: 25–46. [\[CrossRef\]](#)
- Mahoney, James. 2012. The Logic of Process Tracing Tests in the Social Sciences. *Sociological Methods & Research* 41: 570–97.
- Mastrofski, Stephen D., Roger B. Parks, and John D. McCluskey. 2010. Systematic Social Observation in Criminology. In *Handbook of Quantitative Criminology*. Edited by Alex R. Piquero and David Weisburd. New York: Springer, pp. 225–47. [\[CrossRef\]](#)
- Mauss, Marcel. 1923. Essai Sur Le Don Forme Et Raison De L'Échange Dans Les Sociétés Aarchaïques. *L'Année Sociologique* 1: 30–186.
- Mehan, Hugh. 1979. *Learning Lessons, Social Organization in the Classroom*. Boston: Harvard University Press. [\[CrossRef\]](#)
- Mosselman, Floris, Don Weenink, and Marie Rosenkrantz Lindegaard. 2018. Weapons, Body Postures, and the Quest for Dominance in Robberies: A Qualitative Analysis of Video Footage. *Journal of Research in Crime and Delinquency* 55: 3–26. [\[CrossRef\]](#) [\[PubMed\]](#)

- Nassauer, Anne. 2012. *Violence in Demonstrations: A Comparative Analysis of Situational Interaction Dynamics at Social Movement Protests*. Berlin: Humboldt Universität zu Berlin.
- Nassauer, Anne. 2016. From Peaceful Marches to Violent Clashes: A Micro-Situational Analysis. *Social Movement Studies* 15: 515–30. [\[CrossRef\]](#)
- Nassauer, Anne. 2018a. Situational Dynamics and the Emergence of Violence During Protests. *Psychology of Violence* 8: 293–304. [\[CrossRef\]](#)
- Nassauer, Anne. 2018b. How Robberies Succeed or Fail: Analyzing Crime Caught on CCTV. *Journal of Research in Crime and Delinquency* 55: 125–54. [\[CrossRef\]](#)
- Nassauer, Anne, and Nicolas Legewie. 2018. Video Data Analysis: A Methodological Frame for a Novel Research Trend. *Sociological Methods & Research*. [\[CrossRef\]](#)
- Norris, Sigrid. 2004. *Analyzing Multimodal Interaction: A Methodological Framework*. New York and London: Routledge.
- Pauwels, Luc. 2010. Visual Sociology Reframed: An Analytical Synthesis and Discussion of Visual Methods in Social and Cultural Research. *Sociological Methods & Research* 38: 545–81. [\[CrossRef\]](#)
- Philpot, Richard, and Mark Levine. 2016. Street Violence as a Conversation: Using CCTV Footage to Explore the Dynamics of Violent Episodes. Paper presented at the Annual Meeting of the American Society of Criminology, ASC 2016, New Orleans, LA, USA, November 15–19.
- Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California Press.
- Reicher, Stephen D. 2001. The Psychology of Crowd Dynamics. In *Blackwell Handbook of Social Psychology: Group Processes*. Edited by Michael A. Hogg and Scott Tindale. Oxford: John Wiley & Sons, pp. 182–208.
- Reiss, Albert J. 1971. Systematic Observation of Natural Social Phenomena. *Sociological Methodology* 3: 3–33. [\[CrossRef\]](#)
- Salganik, Matthew J. 2017. *Bit by Bit: Social Research in the Digital Age*. Princeton: Princeton University Press.
- Small, Mario Luis. 2009. ‘How Many Cases Do I Need?’ On Science and the Logic of Case Selection in Field-Based Research. *Ethnography* 10: 5–38. [\[CrossRef\]](#)
- Smith, Richard L., Clark Mc Phail, and Robert G. Pickens. 1975. Reactivity to Systematic Observation with Film: A Field Experiment. *Sociometry* 38: 536–550. [\[CrossRef\]](#)
- Strauss, Anselm L., and Juliet M. Corbin. 1998. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*, 2nd ed. Thousand Oaks: Sage.
- Swedberg, Richard. 2011. Theorizing in Sociology and Social Science: Turning to the Context of Discovery. *Theory and Society* 41: 1–40. [\[CrossRef\]](#)
- Swedberg, Richard. 2014. *The Art of Social Theory*. Princeton: Princeton University Press.
- Sytsma, Victoria A., and Eric L. Piza. 2018. Script Analysis of Open-Air Drug Selling: A Systematic Social Observation of CCTV Footage. *Journal of Research in Crime and Delinquency* 55: 78–102. [\[CrossRef\]](#)
- Tilly, Charles, and Sidney G. Tarrow. 2006. *Contentious Politics*. Boulder: Paradigm Publishers.
- Vrij, Aldert, Lorraine Hope, and Ronald P. Fisher. 2014. Eliciting Reliable Information in Investigative Interviews. *Policy Insights from the Behavioral and Brain Sciences* 1: 129–36. [\[CrossRef\]](#)
- Wiles, Rose, Andrew Clark, and Jon Prosser. 2011. Visual Research Ethics at the Crossroads. In *The SAGE Handbook of Visual Research Methods*. Edited by Eric Margolis and Luc Pauwels. Los Angeles: Sage Publications Ltd., pp. 685–706.

