Escalation of Violence in Unclear Situations - A Methodological Proposal for Video Analysis
Reichertz, Jo

Escalation of Violence in Unclear Situations –
A Methodological Proposal for Video Analysis

Jo Reichertz[*]

Abstract: *»Eskalation von Gewalt in unübersichtlichen Situationen – ein methodisch-methodologischer Vorschlag zur Videoanalyse«*. In everyday life, it is rather rare for conflicts to escalate and for violent acts to occur between the parties involved. And when it does, there are usually only a few people involved. Videos of such events are therefore still relatively easy to analyse (despite their complexity) because the events have a centre (monocentric) and the action is sequential, driven only by the interaction dynamics of the participants. However, this looks completely different when the videos show a very large number of people (i.e., everything from 20 people upwards), who belong to different groups with different interests, meet in a specific, pre-structured confined space, and conflicts and violent actions repeatedly arise at different places in the action. Such events often have several and changing centres (polycentric), and there is often an alternation between escalation and relaxation. In addition, several strands of action run parallel to each other and also influence each other (intermediary). Videos of such events pose enormous challenges to social scientists. In my article, based on the analysis of a video capturing the storm of the singling out facility in a soccer stadium, I will show how such complex escalation events can be effectively analysed. The analysis itself consists of a combination of video, interview, and dispositive analysis. The paper will also show why escalation processes cannot be understood and explained solely from what happens in the situation, but the meso and macro levels in which the situation is embedded must always be taken into account as well.

Keywords: Sociology of violence, micro-sociology, hermeneutic video data analysis, escalation of violence.

[*] Jo Reichertz, Institute for Advanced Study in the Humanities (KWI), Essen, Germany; Jo.Reichertz@kwi-nrw.de.

1.   Introduction[1]

It is rather rare for conflicts to escalate to the point of violence in everyday life. However, if this does happen, there are usually only a few people involved. The people involved in conflict are usually easily identified by observers and the succession of their actions are more or less easily recognised. It is relatively simple to analyse videos of such incidents (despite their complexity) because there is only *one* centre of action and attention, and because the action goes on sequentially. It is driven only by the dynamics of the interactions of those involved.

This is different if there are many people in the video (more than 20-25), who belong to diverse groups with different interests meeting in a prestructured small space. Examples for this are regular and ritualised encounters of football fans, in and around stadiums. Part of the plot are police officers, security, fan representatives, fan supervisors, and representatives of different media. Such events are often polycentrical. Just as often, there is an alternation between escalation and relaxation. Additionally, there are several plots affecting each other unfolding at the same time. Furthermore, these actions are greatly prestructured by buildings (e.g., admission facilities), which enable some actions and prevent others.

The action is thus confusing and emotionally demanding for those involved. It shifts the focus on a specific challenge of established procedures: examining situations of interaction and situations of confrontation hermeneutically and based on video (Knoblauch and Tuma 2011; Reichertz and Englert 2021; Tuma 2016; Meier zur Verl and Tuma 2021). These procedures are primarily aimed towards analysing centred encounters with few people. There are scientific (for an overview, see Moritz 2014 and Moritz and Corsten 2018) and non-scientific (see Tuma 2016) methods and tools for transcription, notation, coding, and analysis of video material. However, these systems of transcription and video-interaction are used to analyse scientific field studies (videography) or methods to analyse arranged film (with pre- and postproduction). They are designed (as stated before) for video with *one centred* interaction with a *limited number* of participants, who are involved in *one* plot, which unfolds *sequentially*. To put this more simply, since the acts of interaction of every single person can be focused exactly, the communicative

structure of the observed situation can be reconstructed very well. These procedures are not very suitable if (a) the events in the video material have more than one centre (=multicentric) or if (b) they are continually influencing each other (=interdependent) and are involved in multiple different contexts (=polycontextural).

The aim of my article is to introduce a *method of video analysis* that does the complexity of the event more justice – namely the practice of *story-based data fixing- and analysis strategy.* To achieve this, I used the analysis of a video that shows the rushing of a turnstile in a football stadium. The video shows a complex sequence of escalation. The analysis itself consists of a combination of video, interview, and dispositive analysis to show the events that took place as not only situational or culprit-centred, but rather as a process and imbedded into the trajectory of going to a football match.[2] My article itself is structured as a story that explores how I "discovered" this story-based procedure bit by bit with my team of researchers.

2. Attending a Football Match as Trajectory

Fistfights are predictable because they occur regularly before, during, and after football matches, in- and outside of the stadium. Sometimes they are planned, but they often happen randomly but rarely disorderly. Situations escalating to severe physical violence between the involved groups are a longstanding tradition. This tradition takes its justification from stories and *myths.* There are also specific practices of performing violent acts, which can be described as *rituals* (Turner 1995).

The escalations that happen at the stadium and nearby football matches arises out of the social type of the event itself: two parties engage in physical competition or combat. Both the players on the pitch and the fans (home and away) engage in some sort of battle. These battles among players and fans go

---

[2] If social-scientific literature on the topic of emergence of physical violence is viewed, it becomes clear that there has been a shift in explanatory patterns. This shift can be roughly described as a shift from culprit-centred explanatory approaches to more situation-centred explanatory approaches (see Hoebl and Knöbl 2019 and the introduction to this book for more detail). The *culprit-centred* perspective highlights the social-structural or individual factors and positions that supposedly are present for violent people (von Trotha 1997). In contrast, the *situational* perspective argues that even registered offenders do not act out violence most of the time. According to this, violence arises out of the situation and only exists in situations, which is why these situations need to be examined (see Collins 2011; Equit, Groenemeyer, and Schmidt 2016). The situational perspective has recently gotten criticised for being too narrow if it only observes the situation, which led to the view being shifted to the whole *process* of the escalation of violence over the years; the loosely coupled chains of interaction of different actors, who, in the end, but not necessarily, lead to violent actions (for more detail, see Reichertz and Keysers 2018; Reichertz 2018a; Hoebel and Malthaner 2019; Hoebl and Knöbl 2019; critically to this Kühl 2021). This article follows this lead.

back decades, which is why characteristic frameworks, practices, traditions, and routines have formed. Knowing about these traditions is vital for someone trying to understand and work an event like this. Even if the occurrences are unique, they are re-enactments of a known type of occurrence but differ in specifics.

To put this in the words of logic, every single specimen of an event like this is a *token* of a specific *type*, which is why the specifics of a token are only visible in the context of the type. Attending a football match is a regular occurrence, where many people act independently. If an attempt to detect the most important "players" involved at a football game is made, it is easy to see different groups of interest. Police and security, for example, are looking to keep or restore order, while some fans want to overrule. For example, some Ultras3 have expressed plans to perform a large pyrotechnics act, which security aims to prevent at the gates. Because these goals (partially) contradict each other, there is no cooperation (in the narrow sense of the word) among those involved, they base their actions on the actions of the other party. To be able to achieve this, the parties watch each other and make their actions visible (sometimes unintentionally because they are attempting to hide parts of their actions). They *coordinate*. From a sociological standpoint they *play* (in the sense of Goffman 2005) together, they *collude*.[3]

There is something that makes this form of coordinated acting special: the actors adjust their actions to each other, but there is nobody tasked with reaching the set goal. There is no one person responsible for the target orientated chaining of single actions. The process as a whole exceeds the planning horizon of those involved. For social processes that are not performed by *one* subject but are achieved by independent and rivalling efforts of several people, Anslem Strauss has suggested the term *trajectory* (1991, 1993; see Soeffner 1991; Reichertz 2016a).

Trajectories are processes. Their course of events can be divided into stages, which are aimed toward a goal, without it being known whether that goal will be reached or not. To achieve said goal, the course of events must take a specific path. The path is not predetermined, there are always alternatives, which is why the concrete design of the process is always subject to the communicative reassurance among the people involved. This process of permanent communicative reassurance and construction builds the course of events – it is thus the result of communication and interaction.

---

[3] Ultras are organized groups of fans. They are much more committed to their club than the rest of the fans: they dress in the club's colours and cheer loudly for their team during the game. Sometimes they also set off fireworks, which is forbidden in Germany. Violent confrontations between the ultras of the soccer clubs are also rare.

[3] See Laing 1977. According to this, collusion (lat. colludere = play together) is a pattern of interaction among two or more participants that functions like a cooperation, since it consistently matches the acting of those involved. This usually happens unconsciously.

3.   Involved Actors

There are different actors involved in the trajectory of going to a football match, who get involved into the action at different times in different ways (Reichertz 2018a, 260-9). The phase of visiting fans arriving at the gates and the ensuing inspection is important and delicate. The fans are checked for entry tickets and if they brought prohibited items to the stadium (e.g., flares). The reason this task is so delicate is that to achieve this, the large group of fans needs to be broken down into singular people searched for contraband. This is made even more difficult if the fans brought legal material to support their team (e.g., drums, backpacks, flags, etc.) and used it to smuggle prohibited items.

  The most important *personal* actors at this event are the *visiting fans* and especially the *Ultras* organised within the club. There are also fan representatives of the clubs themselves present at all times to prevent misunderstandings and turmoil; they are available for communication. On the other side are *private security workers* working for the stadium operators, who are supposed to search for prohibited items. There is also always a large number of stand-by police officers who are supposed to restore order in case of violence. Additionally, there are always *police officers familiar with the scene* and *people in charge* from the stadium operators among those involved. They are supposed to ensure that things go peacefully and conflicts are solved in a communicative fashion. Of course, there are always several people working for various media outlets present who record and publish images, sounds, and quotes.

4.   Polycontexturality and Complexity

Everything we perceive that others express, or everything we *see* or *hear* that others do, is understood as communication. This is being used to adjust our own acting to the acting of others (detailed: Goffman 1974; Reichertz 2009). Therefore, it is always important to determine the communication space when conducting social-scientific analysis. In the evaluated case, the situation is rather unclear. This is not only because there are so many people communicating over each other at the same time, but also because of loud yelling, whistling, cries of pain, protesting, and insulting. All these are expressions everyone present has access to.

  Hence there are many direct and indirect, explicit and implicit, and institutionalised and informal relationships of communication. There is a wide variety of direct and digital communication *within* the different groups (ultras, police, security, stewards, media).

There is far less communication *between* the groups, some groups (at least officially) do not communicate at all (police – ultras).[4] Because such firm and institutionalised ways of communication between security (police and private) and ultras or between stadium operators and ultras are absent, police and stadium operators use powerful loudspeakers to make pleas or give instructions.

Another possible way of communicating between the groups is informal communication on-site in the given situation. This, however, can only develop communicational power if people are present who are known in the scene and who have achieved *communicational power* (Reichertz 2009) with the other groups. This means these people must have some history together, during which they have earned the others' trust; clubs' fan-representatives and scene-savvy officers on one side, and individual experienced fans on the other can diffuse dangerous situations before they happen. However, there is no reliable or tenable institutionalised communication between fans/ultras and security forces.

The situation is made unclear and confusing for those involved or looking on by the various means of communication. Since all participants turn towards selected acts of expression and use said acts as a cause for their own communication, their own actions are influenced by different contexts. Then again, their own actions are interpreted by others in different contexts, who then turn it into a subject of communication themselves. Hence, because these contexts of communication can be so diverse, it is not observable who acts when and upon what. Such situations are thus structurally marked by *polycontexturality*.

5. The Analysed Case

The base of my example analysis is a YouTube-video from 2017, titled *Polizei prügelt auf VfB-Stuttgart-Fans ein* (*Police beat up VfB-Stuttgart-Fans*).[5] The video shows violent conflicts prior to a Bundesliga match between VfB Stuttgart and Arminia Bielefeld. The match took place in Bielefeld. During admittance checks, conflicts arose between private security forces from Arminia Bielefeld, fans from Stuttgart, and later, between police officers and the same fans. The video was uploaded to YouTube during the night of the game, on 17 April 2017, by a dedicated member of Stuttgart's fan scene. It has remained online until early 2019.

The video was taken with a mobile phone and is 12 minutes and 31 seconds long. The camera orientation, for the most part, does not change. The camera

---

4   Supposedly, it is possible (according to scene-savvy police officers) for officers to contact ultras via a lawyer in extreme emergencies in some cities.
5   See https://dx.doi.org/10.14279/depositonce-15993.

records (sometimes pivoting and focusing events in other areas) an open area (about 30 m wide, 20 m deep), where a metal admittance gate is situated. The video shows a situation during admittance before the match from the side. The distance from the camera to the gate is about 30 m (see image 1). Since the available data is viewed from a *medium* distance over a *longer period of time*, there are four methodical and methodological consequences. The result shows that data cannot be theoretically innocent, it rather suggests and enables certain views and theoretical findings.

Firstly, although there is much to hear in the video, it is mostly undifferentiated noise consisting of yelling and whistling, among other sounds. The people present participate vividly in what is happening; they seemingly insult and threaten, but also appease and calm others down. All that, however, blends into a mess of sound. This fact, that the verbal interactions are not distinguishable, is important for the analysis. It is neither possible to make out what exactly is being said, nor is it possible to point out who says what when to whom, and with what consequences. This results in having to analyse such a video entirely without spoken language. This prevents the analysis of video to turn into an analysis of spoken material, and thus, an analysis of text.

Secondly, the events shown in the video are not fixated on a micro or nano level of communication. Thus, the only actions available are those shown from a *medium distance*. This is why gestures and facial expressions are not well visible, which, in turn, is why they will not be analysed. Thus, an advantage that was hoped for, especially when analysing video, is lost. The acting and moving of the bodies and the moving of hands and feet are what is mainly identifiable in the video.

Thirdly, because of the distance, *singular* actors are not focused on (the argument between two people with an ensuing fight), but rather the whole situation. Because of the stable camera, the attention is evenly distributed among the many people in the video. It is hence visible where people are, where focal points of action are taking place, and how they are distributed across the space.

Fourthly, there is not just one or two minutes of action being shown (like for example the raid of a gas station register), but rather actions being carried out by different people in different places that influence each other in a span of about 12 minutes. All these actions take place in a shared space of perception. It is thus not only visible where people and focal points of action are and how they spread out, but the view also shifts toward the timing of action in different spots and their interdependence. The video is not a picture puzzle. There are many people and things in picture puzzles, but the interactions of the people are centred and secluded from each other. Thus, they do not influence each other.

Here, it is shown that data cannot be theoretically "innocent," data always shows a certain interpretation of an event and allows theoretical insights to happen. If one focuses on interactions between two or three people, one can only see their share of how the action evolves. However, if one shows the crowd- and fringe conditions, as well as how things were before and how they are afterwards, then reconstructing what has happened shows different results rather quickly.

Image 1 Screenshot of the Analysed Video

The title reads "Police beat up VfB Stuttgart fans + pepper spray – Bielefeld vs VfB Stuttgart 2:3 17.04.17.

6.   A First Try of Data-Fixation in Three Steps

The video shows how, on Monday night, during admission of Stuttgart supporters, there were significant scuffles between people who travelled to watch a football game (security workers and police). This takes place in front of the stadium in Bielefeld. During the video, numerous typical processes of escalation (compression, synchronisation) and characteristic forms of rampage at football games (e.g., throwing of objects, storming the admittance gates) are visible (Reichertz 2018a, 288-91). In total, we made out nine different incidents where, at different places and times, altercations happened.

Sometimes there were many people involved, there was punching and kicking, and the police made use of pepper spray.

The conflicts arise concretely in front of the first gate of admission (separating the visitors to be searched individually), as well as within the gate (searching for dangerous goods) and at the exit of the gate (entrance to the stadium). There was a total of about 150 to 200 people across all the groups mentioned earlier involved. Out of those, about 70 to 80 were actively engaging in violent acts, some more than others. Even if the others were not actively involved in the violence, none of them were mere bystanders; everyone was involved in the situation, there was a feeling of great tension. Some among those not involved proceeded to push from behind, others verbally provoked, others took photographs, while others tried to deescalate the situation.

Everyone present was relevant for what was going on because they actively contributed to the specific form of the event. This means, for the purpose of this analysis, the conduct of every person needs to be identified and tracked over the duration of the video. Therefore, there needs to be a process to isolate the actors and to watch their conduct in the course of time. We proceeded in three steps.

6.1   Construction of a Spatial Grid

The first step was constructing a spatial grid in the frame to describe location and movement of any given person. To achieve this, we divided the space visible in the video into different corridors and zones. We then layered said grid over the individual frames (see image 2). Corridor 0 is the square in front of the stadium. Corridors 1 to 4 are the passageways between the admission gates, corridors 5 and 6 are within stadium grounds.

Image 2 Division into Corridors for Split Up Fixation of Data
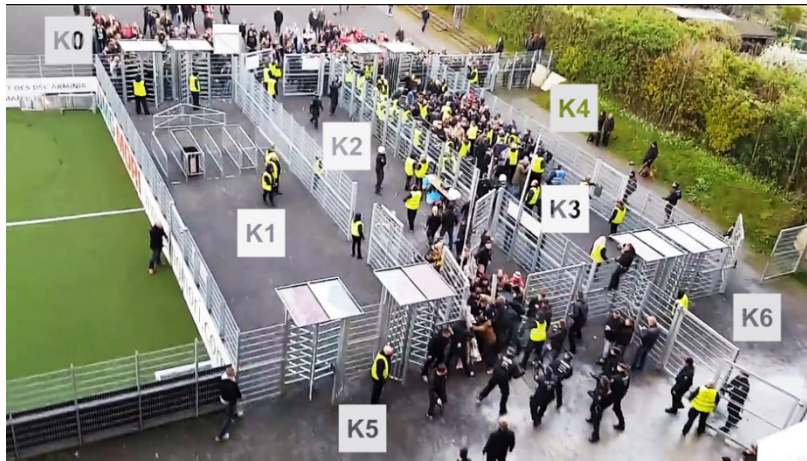
The screenshot shows (among other things) how a group of fans is trying to open the gate into corridor 2 forcefully. Police officers are trying to hinder them by using pepper spray.

6.2   Short Description of the Video, Oriented Towards the Progress of the Video

The second step was a short description of the video to add a perspective of progression to the segmentation. As the video starts, loud noise is audible: voices, yelling, barking dogs, bustling people; especially in corridor 3. There is also a lot of movement in corridor 2. There are fans in both corridors, with corridors 5 and 6 not in use at this time. Many fans are still in the courtyard outside the stadium and have yet to go through the turnstiles, while others are already inside the admission facility to be searched. Meanwhile, some other fans are already on stadium premises. At the start of the video, a fist-fight is happening in corridor 3, involving security and fans. After a short time, more security and some police officers get involved. Shortly after that, a second group of fans starts some activity in corridor 2. The group consists of about 25 people, who are obviously being denied immediate access to the stadium premises. When the door to the stadium premises has to be opened for a seemingly injured person led from corridor 3 through corridor 2 by a woman and security personnel, said group of fans in corridor 2 seizes the opportunity and tries to get into corridor 5, and thus onto stadium premises. A group of young men, who are already in corridor 5, supports this by storming the gate to prevent the woman from closing it behind herself and the injured person. Some fans manage to step into corridor 5; however, the whole group is forced back into corridor 2 by police officers using pepper spray.

   A minute later, people in corridor 2 try to force the gate open once again. This time, however, fans who are already on stadium grounds help as well.

Once more, police prevent this from happening with pepper spray. Fan representatives then try to deescalate the situation in corridor 2. As a result, some fans are led from corridor 2 to corridor 3. From there, they are led through a turnstile into corridor 6. Soon thereafter, a scuffle between some fans and one member of security breaks out. This is cooled down by a police officer spraying pepper spray through the fencing between corridor 5 and 6. This leads to another confrontation in corridor 6, since some fans, who are already on their way into the stadium, feel provoked by the pepper spray being used against them once again. At the same time, a formation of police officers is led into corridor 6 via corridor 4. They separate the two sides but add more fuel to the fire at the same time. Meanwhile, more officers arrive in the courtyard via squad car. The video ends here.

6.3   Transcription

The separation of the admittance facility into six corridors and the following short description served the detail-orientated transcription (see image 2) in step three. The separation of the image into segments was done in accordance with the areas of action opened and closed for the actors by the admittance facility. Thus, we followed the logic of the field, the literally predefined paths and boundaries. Since it did not seem possible for one person to transcribe everything that happened in the video at the same time, we split the work up among us; every member of the project staff was assigned a corridor and transcribed the events that happened in the allotted corridor according to a specific timeline.

To be able to transcribe, staff watched the video not only once, but hundreds of times on repeat, slower or faster, with or without sounds, zoomed in or even with different (colour- and sound-) filters, to be able to assess specific aspects better or differently. All this was done to adequately capture and fixate what was happening in the video material. In the end, we fused the developed transcription of the singular corridors into one sheet so that all the simultaneously ongoing events can be "read" like sheet music.[6]

7.   Limitations of the Developed Procedure

Separating the images and splitting the workload made the transcription doable, however, it did not do what it was supposed to: to reduce the complexity of the recorded events. The sheet itself ended up so complex it was barely readable. To be clear: for interpretating video, it was helpful to transcribe the

---

[6] The transcription of the video was achieved using the software *Feldpartitur* (Moritz 2011). The utilised categorisation of the notation was not given from the start, rather it emerged from the transcription and the work associated with it.

material in a detailed fashion, since different people may see more things, or they see them differently.

Transcription of video does not consist of only one perception, but always and necessarily of interpretation (Reichertz 2016b; Moritz 2011). The data is thus not a result of simple fixation of something observed, but a result of interpretation. This is, by the way, why it is of such importance for the transcription work to remain within the research project instead of being sent off to an external institute. Someone who works with externally created transcriptions knows neither their own (raw) material, nor the date of analysis (transcript, sheet, etc.) obtained out of it sufficiently. For scientific interpretation of video material, it is vital and obligatory to adopt the data as one's own with exceptional intensity.

However, transcribing what is perceived to make this process of research available and comprehensible for the own team or for others is only one purpose fixation of data. In this sense, video transcriptions serve the scientific obligation to produce proof and the purpose of communicating within the research team. The *actual* purpose of transcription, however, is to create a product, which can be used practically for further analysis. This necessarily entails handleable data, which reduces the complexity of the video but *conserves* or *condenses* the analytically important aspects being produced. Only then can it enrich and accelerate video analytical work.

We did not really manage such conservation – on the contrary: we had 20 pages of a video sheet that only shows the notation regarding what *one* person did over a timespan of 12 seconds. The attempt to simplify the video through notation led to the data-fixation being so complex, it almost became obsolete, since the context of the action was now barely recognisable.

The sheet as a whole proved to be extensive and not very helpful during hermeneutic interpretation. In this context, they cannot be compared with detailed verbal transcriptions of interviews – at least not regarding the significance a transcription can have in the context of hermeneutic interpretation. The non-handleable lack of clarity in the fixation of data once again bears the question: How can multicentrical video interactions be fixated sensibly? Our answer to this question is not to trash the whole sheet, but to analyse again and again – shifting from looking at *all events* to *focusing on separate people* who were of particular importance for the process by significantly shaping the events. By looking at people and their actions over a certain length of time, we managed to write a story, to isolate a chronological sequence of succeeding action, and to (re)construct their significance hermeneutically.

8. A Story-Based Strategy for Fixation of Data and Analysis for Complex, Multicentrical Chains of Events

We knew from the start that we had navigated into difficult and mine-laden terrain with the decision to write *stories* of consecutive actions. The term "story" has been *en vogue* for some years, which did not really help its clarification (Gottschall 2013). If we have chosen to try this path nevertheless, it is mostly out of hope of being able to comprehend video-data and their meaning in more depth – mindful of the fact of having to consider the stories' character of construction when finally analysing.

If the situation is confusing and polyphonic, such stories help create order. The polyphony of an examined sequence of action is revealed through many stories, which arrange contrasting courses of acting along a timeline. More importantly, this is about making data visible for interpretation, aided by these stories. In no case are narrative structures imposed on the diversity of the world – neither central fix- or turning points are being introduced nor is there grading of what is told in view of the cause of the telling. It must not be overlooked (this needs to continually be reflected upon during analysis) that these reconstructions of consecutive actions in the form of stories are *constructions* that result out of interpretation and deduction of the interpreting person. Thus, the expatiating of such stories is, without a doubt, a communicative construction (Reichertz 2009; Keller, Knoblauch, and Reichertz 2012; Reichertz and Tuma 2017). With this practice of writing stories, the writers always and necessarily make use of the arsenal of the societally available narrative styles and figures. This must not be forgotten during interpretation later, either.

Writing a story means, in a *practical* sense, to look at one or several acting individuals (this works best with two or three observers) and to isolate their respective movement and their results and to (re)constitute their social meaning. This hermeneutic procedure "translates" the isolated body movements into actions and views these actions as a solution to recent problems of acting, or as responses that result from the (further) acting of the actors. Stories were searched for until there were no empty spaces left in the footage.

In the following passages, I will play back *three* of these stories of consecutive actions. They are to serve as an example to illustrate the procedure used when fixating and analysing data. They are also used to show how answers and questions arose out of the explication of stories. These answers and questions guided the purposeful sampling of further data and led to *thick descriptions* (Geertz 1987) and different formations of theory of case overlapping significance via *triangulation* of *perspectives, data,* and *methods*. For ease of addressing the individuals during our evaluation sessions, we retrospectively gave them a name that signifies a *special trait* or *function*.

8.1   The Arsonist

The arsonist is a very conspicuous person. He is about 35 years old, has ash blonde hair, not very tall, sturdy, and easily recognised as a member of the stadium's security brigade due to the neon yellow vest he is wearing. The first time he appears in the video is in the beginning, in corridor 0, so *in front of* the admission facility. There, he helps to protect the access to the gate to corridor 2 from being stormed by the fans. Corridors 2 and 4 are only accessible to security personnel, so they can reach the other corridors quickly, if need be, without having to pass through the turnstiles. Corridors 2 and 4 simultaneously serve as emergency exits, which is why the gates may not be locked. Therefore, they are being guarded separately. Obviously, the arsonist has seen a conflict arise with fans in corridor 3. These fans had already passed the turnstile into corridor 3. First, insults are exchanged, then a scuffle ensues. This causes the arsonist (and later police officers and more security) to open the gate from corridor 0 to corridor 2 and hurry to the already open gate between corridor 2 and 3.

After he got into corridor 3, he aggressively pushes his way through the fans and stewards towards the turnstile, where the scuffle between 6-7 fans and 4 private security members is going on. He purposefully gets involved by punching a fan with his fists. This leads to the group of fans to dissolve and the scuffle to end.

During the next few minutes, the arsonist, still aggressively pushing and shoving, makes his way towards the gate situated between corridor 3 and 4. He punches and kicks fans repeatedly. He pushes one man, who was involved in the scuffle, into the arms of four police officers at the gate, who arrest the young man and lead him into corridor 4. Then, the arsonist moves close to the turnstile that leads to corridor 6. From there, he watches, prancing tensely, until two attempts to open the gate to the stadium premises violently by fans in corridor 2 are made. As the situation calms down and the fans are being diverted to corridor 6 via the turnstile, the arsonist passes the turnstile as well. Immediately after he has passed the turnstile, he attacks a fan who had just gotten there. This causes a bigger violent altercation between fans and security personnel. Newly arrived officers in corridor 4 defuse the situation, which took a considerable amount of officers.

The arsonist is the most mobile actor on-site and commits violent acts in all corridors that contain fans. He has no firmly assigned spot and obviously has the liberty to access all corridors. He traverses the facility, obviously intending to track down especially active fans and to attack them with his fists and feet. He is looking for confrontation. He is actively involved in six short brawls. He does not calm situations of conflict down, he creates them. Since he can move freely through the entire facility, he can create turmoil at every place, which is exactly what he does undisturbedly.

8.2   The (Fan) Reporter

A man about 20 years old in a dark windbreaker is notable because he is standing in an elevated position. He achieved this by climbing up the fencing between corridor 2 and 5 (from inside the stadium premises) like the steps of a ladder to get a better vantage point. He is holding a camera up with a straight arm to record the fight behind the turnstile in corridor 3, the one the arsonist is involved in. As the Stuttgart fans try to storm the gate in corridor 2, he records the ensuing altercation between fans and police. As soon as this ends, a brawl breaks out directly in front of him, in corridor 3, which he records as well. As a police officer attacks a fan behind him, the (fan) reporter turns around to capture this event on video. Then a larger fight between fans and security, which the arsonist is involved in, breaks out directly behind him in corridor 6. The young man goes to sit down on top of the fencing that separates corridor 3 from the stadium premises and records the events in corridor 6. After a short while, a steward approaches and orders the (fan) reporter to get off the fencing. He obliges the request and leaves the frame in direction of the stadium. In total, the young man sat on top of the fencing for about 5 minutes. He mainly recorded the violent altercations that broke out in different locations. It is apparent that he focused the camera mainly on stewards, security, and police officers.

8.3   The One Striving for Order

The one striving for order is an approximately 40-year-old woman with long, blonde hair who works for the home club. It is her job to take care of the visiting fans. For this reason, she is free to move between all corridors. At the beginning of the video, she is in the middle of corridor 2. She accompanies two fans, who apparently got injured during the confrontation in corridor 3, into the stadium grounds. For this reason, they must move through corridor 2 to get to the unlocked door to corridor 5. The woman leads the formation, opens the gate between corridors 2 and 5, lets the injured pass and then follows them, closing the gate behind her. A small group of fans that is being searched in corridor 2 tries to keep the gate open to follow her into the stadium. The woman tries to shut the gate with more force. In this moment, a group of fans from corridor 5, within stadium premises, storms against the gate from the inside, violently pushing the woman into the now open gate. She hangs on to the gate's handle while some fans pour into the stadium. She continues hanging on, even as the police push the fans back into the corridor using tear gas. She goes through the now clear gate, past the fans, who are suffering the consequences of having been tear-gassed, through the gate, pulls it shut behind her and leaves the scene without saying a word.

We "detected"/constructed about 30 to 40 stories in the actual empirical analysis. Surely, it would have been possible to detect or construct many

more stories, as many as there were actors involved. The admission facility itself was also involved – very dominantly so (see Reichertz 2018b). Therefore, there is also a story about the admission facility. This is a story that has to be told, without me making it into an actor (for more detail, see Reichertz and Keysers 2018).

8.4 The Admission Facility as an Imperative of Steel

Even though the explication of stories made much of *what* happened in and around the admission facility clear, it did not explain *why* it happened to me. Since so many actors resorted to similar deeds (scuffles, punching, kicking) at different times, the question, whether the entire setting of events, literally the frame (the admission facility), was greatly involved in causing these events.

This led to us, the project staff, to take a closer look at the admission facility, which we understood as a facility for searching and singling out. After all, it shapes the trajectory of going to a football match fundamentally – by granting or denying access – and condenses thus collective experiences, which we (re)constructed via individuating stories.

A turnstile, at its core, is an *artefact*. In the sense of Foucault (1978), it can be viewed as a contemporary, preventive disciplining-*dispositive* for handling certain groups at big events that implies effects on the subjectivity of everyone involved (see Reichertz and Keysers 2019; Reichertz 2018a) – as man-made *imperatives from the past* that open or close options of acting (Reichertz 2018b). Artefacts open paths that were created by earlier societies and are supposed to make continuing the same way possible. On the other hand, they erect prohibition signs or even close some doors. Still, artefacts or dispositives do not determine particular acts, they rather make some paths easier and others harder. They sanction certain paths, sometimes negatively, sometimes positively, and sometimes they erect walls and fences that are difficult to surmount.

The entrance into the guest fan area in Bielefeld is an about 20 metre long and 12 metre wide, sectioned *isolation system*, a security related artefact made from taller than head high steel beams. Its purpose is to steer the admission to the area behind it into organised lanes. The isolation system consists of four parallel corridors, two of which are corridors intended for guests that can be entered and exited via turnstiles on either side. Between and next to those guest corridors are two corridors, which are not secured by turnstiles. Access to these is only granted to stadium personnel and police, barring emergencies.

The visitors' corridors are accessible via turnstiles of steel, which only let one person through at a time; furthermore, they can be locked at any time. To maintain the singling out within the corridor, there is a so-called *barricade*,

a construction made from steel. This *barricade* is there to create paths within the corridors, allowing security personnel to search visitors quickly. There is another set of turnstiles at the end of the corridor that needs to be passed through to gain access to the stadium (see image 3).

Image 3   Scheme of the admission facility in Bielefeld
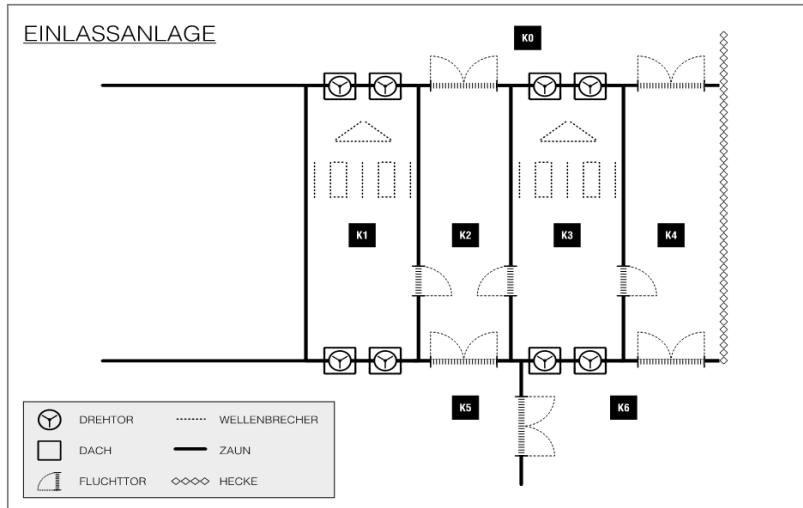
Legend: K(number) = corridor; Drehtor = turnstile; Dach = roof; Fluchttor = emergency gate; Wellenbrecher = barricade; Zaun = fence; Hecke = hedge

There are (unlocked) gates between the corridors, which actors with appropriate clearance can use to move from one corridor to the next. This is unavailable to those who do not have this clearance.

This construct of lattices cannot ensure admission in an appropriate and orderly manner on its own, however, it depends on human actors (security and stewards) to be "maintained" and utilised. These actors, who are "installed" into the facility on match days to operate it have plans, laws, policies, training, and experience to aid them when operating the facility. Sometimes, they must interpret them anew, though. The iron construct and the people integrated with the lattices create a unit; a mechanism of social selection, which replies to a special historical problem with German Bundesliga matches.

Even though everyone involved in these processes talks, or, rather, shouts, the facility as a whole is not designed for organising conversation or debate. Its goal is organising the *bodies* of the people involved. It is not about the better argument, it is about assertiveness, which, in the end, depends on having more *physical strength*. Not only does the facility need those operating within it, those operating need the facility as well. The steel displays an extension of

the operators' bodies. This accompanies an unspoken understanding, that subjects (both fans and workers) are mainly viewed and evaluated by whether they are strong or weak, ready to use violence or not, and whether they are letting the controlling device discipline them or not. *This makes the facility a human and steel device that communicates consistently, albeit silently*: "You, who come to this device, are potentially dangerous, since some people want to do damage in the stadium, or because they want to take dangerous and prohibited items into the stadium. This is why we will slow you down, single you out and inspect you, and, if need, be, detain and deny you admission against your will."

The reply given by Stuttgart supporters to this sort of *preventive pacification* may not be the one the organisers of the event had hoped for, but it is one that follows the line of the subjectivities that were communicated through the social mechanism. This is how the pacification led the chaos and violence that was visible in Bielefeld on that day.

But there was a particularity that day, as we learned during our interviews with actors involved that day (fan advisors and scene-savvy police officers). As shown above, corridor 2 is not accessible to fans due to the gates. However, on that day, stewards and security determined they had no appropriate location to search and inspect Stuttgart's fans and their belongings, and that they had nowhere to collect confiscated goods. They then decided ad hoc to utilise corridor 2 to separate fans with material to be searched. For this reason, a container with a blue bag was installed to collect confiscated material. This is why fans were led into corridor 2 and got the chance to storm the gate in the first place.

Image 4 The Admission Facility in the Visitor's Area of the Stadium in Bielefeld, 17.04.2017
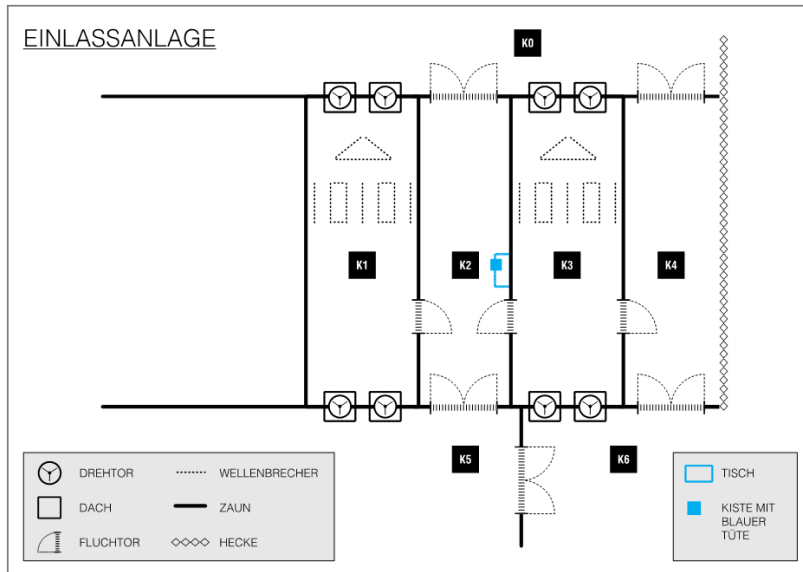
Modifications done by personnel is visible in corridor 2. Tisch = table; Kiste mit blauer Tüte = bin with blue bag.

However, the events that took place that day cannot be explained only on account of the alterations to the facility, which were done with the knowledge of scene-savvy police officers. The installation of a box with a bag and the redirection of the fans via corridor 2 were an exception, which created the opportunity that led to violence. This violence would not have been possible on other days. The reconstruction of what occurred shows that there is a distinctive feature built into the facility that structurally aids the formation of such conflicts. The referred feature are the roofs above the turnstiles, which prevent broad and long objects (drums, flags, poles, etc.) to be taken through the turnstiles. Instead, either the gates into corridors 2 or 4 need to be opened or the objects need to be lifted over the fencing.

If these doors were to remain closed, however, the result is a fierce debate about how the material should regularly enter the stadium. Admission into the stadium will be hindered, discussions that turn into conflicts arise and, especially due to time pressure, escalate into scuffles or violent outbreaks. Drastically, the facility that is supposed to single out people for a more rapid flow of people hinders exactly that because it simultaneously creates a problem, which can easily lead to violence. It solves one problem but creates another. Installing this facility has thus created a new problem, which is here to stay. Such a case is shown in the video we analysed.

9. Conclusion

The goal of my contribution was to show how, during the research process, we produced, organised, and analysed various data. Furthermore, I wanted to show how we created a new practice of data fixation and analysis out of the analysis of a video that did not allow for conventional procedures of analysis. This new practice had to live up to the complexity and polycontextuality of what had happened and to the demands of our hermeneutic-knowledge sociological theory of acting. This is how we dissolved the complexity of the events that took place simultaneously in different locations into actions of different actors and made them visible, available, and thus, interpretable.[7]

At the same time, we gathered more data: we conducted interviews with participants of this event and experts and evaluated documents about this event. The principal purpose of this analysis was to interpret the admission situation as a deed-structuring dispositive. This analysis unearthed that the individual conflicts visible in the video are largely intertwined with the altercation and acquisition of the admission facility by the people involved (both security and fans): the instructions of movement made of steel, which are materialised in the admission facility, did not only permanently solve the problem of singling out, but created many repeating problems, like entering the facility with bulky material.

This is why there is no easy answer to the question why it came to violence on that day in Bielefeld. Factors like time pressure, the loaded emotions of fans before the game, the aggressiveness of certain people (like the arsonist), the atmosphere on site, possible prejudices towards stewards and police officers, and many more played a role. The box for storing confiscated goods installed that day in corridor 2 certainly played a role as well. Just as relevant, however, is the facility itself structurally producing conflicts that could lead to violence, especially in such tense situations. Because this is how it is, processes of escalation in such situations cannot be explained by the conduct of

---

[7] The practice of video analysis introduced here is, without a doubt, not entirely new; it continues earlier works about video analysis (Reichertz and Englert 2021[2013]). Additionally, many things, which were discussed in recent decades in literature about video-analysis and violence-research, are utilized (Collins 2011; Hoebl and Knöbl 2019; Tuma 2016; Nassauer and Legewie 2018). For the practice of video-analysis introduced here, the attention to direct and indirect chains of interaction that lie before the filmed situation, the regard to the sequentially arranged (medially and directly mediated) process of communication within the situation, and the intentions, hopes, and fears aimed towards the situations after the analysed situation (see Abbot 2019) are all central aspects. This is why the similarities between all forms of knowledge-sociologic video- and interaction analysis, which work sequence-analytical or process-analytical, are not pure coincidences.

the actors within the situation, they can only be understood appropriately with due regard to trans-situational events.[8]

In short, even though the evaluation of the entirety of the material hint towards the great importance of communicative processes of togetherness and division on site, the reconstruction of the escalation process in the admission facility in Bielefeld shows that actions of both the past and the future are being carried into the situation and, thus, influence actions. These communicative processes highlight the relevance of (inter)corporal forms of expression on a micro and nano level. For this reason, social-scientific analyses that aim to explain such events only out of the situation always fall short, solely due to their field of vision being too limited. The reason it is limited is the fact that they limit themselves to what is visible in the video. They are blind when it comes to pasts and futures and risk providing insufficiently complex results.

Data Availability

The video data used in this study are available at aviDa (https://fdz-avida.tu-berlin.de), the research data centre for audio-visual data of empirical qualitative social research, hosted by Technische Universität Berlin, and can be accessed here: https://dx.doi.org/10.14279/depositonce-15993.

References

Abbot, Andrew. 2019. *Prozessuales Denken*. Hamburg: Hamburger Edition.
Collins, Randall. 2011. *Dynamik der Gewalt.* Hamburg: Hamburger Edition.
Equit, Claudia, Axel Groenemeyer, and Holger Schmidt, eds. 2016. *Situationen der Gewalt.* Weinheim: Juventa.
Foucault, Michel. 1978. *Dispositive der Macht. Über Sexualität, Wissen und Wahrheit.* Berlin: Merve.
Geertz, Clifford. 1987. *Dichte Beschreibung*. Frankfurt am Main: Suhrkamp.
Goffman, Ervin. 1974. *Frame analysis: An essay on the organization of experience.* Boston: Harvard University Press.
Goffman, Ervin. 2005. *Interaction ritual: essays on face-to-face interaction.* New Brunswick: Aldine Transaction.
Gottschall, Jonathan. 2013. *The Storytelling Animal: How Stories Make Us Human.* New York: Marina Books.

---

[8] This result is, without a doubt, a largely traceable data collection for the project, which poses the question whether we mainly produced data that make our theoretical premises visible. Fundamentally, the question "Do researchers or certain practices of social research mainly gather data that supports their presupposition and thus create a self-fulfilling prophecy?" can be asked. In any case, it becomes visible that media of gathering data carry a certain implicit epistemology of which factors are relevant for recognising certain processes.

Hoebel, Thomas, and Wolfgang Knöbl. 2019. *Gewalt erklären! Plädoyer für eine erklärende Prozesssoziologie.* Hamburg: Hamburger Edition.

Hoebel, Thomas and Stefan Malthaner, eds. 2019. Im Brennglas der Situation. *Mittelweg,* April 1-2.

Keller, Reiner, Hubert Knoblauch, and Jo Reichertz. ed. 2012. *Kommunikativer Konstruktivismus*. Wiesbaden: VS Verlag.

Keysers, Verena, Joana Meißner, Jo Reichertz, and Nils Spiekermann. 2019. Die Einlassanlage von Fußballstadien als Stahl gewordener Imperativ. In *Problemgruppen in Staat und Gesellschaft*, ed. Dörte Negnal, 167-192. Wiesbaden: Springer VS.

Knoblauch, Hubert, and René Tuma. 2011. Videography: An interpretative approach to video-recorded micro-social interaction. In *The SAGE handbook of Visual Research Methods*, ed. Eric Margolis and Luc Pawels, 419–430. Los Angeles et al.

Kühl, Stephan. 2021. Mal wieder ein neuer ‚Turn‘ in der Gewaltforschung. In *Berliner Journal für Soziologie*. https://link.springer.com/content/pdf/10.1007/s11609-021-00448-4.pdf.

Laing, Ronald D. 1977. *Das Selbst und die Anderen*. Reinbek bei Hamburg: Rowohlt.

Meier zu Verl, Christian, and René Tuma. 2021. Video Analysis and Ethnographic Knowledge: An Empirical Study of Video Analysis Practices. *Journal of Contemporary Ethnography* 50(1): 120-144.

Moritz, Christine. 2011. *Die Feldpartitur*. Wiesbaden: VS Verlag.

Moritz, Christine, ed. 2014. *Transkription von Video- und Filmdaten in der Qualitativen Sozialforschung. Annäherungen an einen komplexen Datentypus.* Wiesbaden: VS-Verlag.

Moritz, Christin, and Michael Corsten, eds. 2018. *Handbuch Qualitative Videoanalyse*. Wiesbaden: Springer.

Nassauer, Anne, and Nicolas Legewie. 2018. Video Data Analysis: A Methodological Frame for a Novel Research Trend. *Sociological Methods & Research*.

Reichertz, Jo. 2009. *Kommunikationsmacht*. Wiesbaden: Springer VS.

Reichertz, Jo. 2016a. Weshalb und wozu braucht man den ‚korporierten Akteur‘? In *CSI. Rechtsmedizin. Mitternachtsforensik*, ed. Carina Englert and Jo Reichertz, 149–168. Wiesbaden: Springer VS.

Reichertz, Jo. 2016b. *Qualitative Sozialforschung – Eine Einladung*. Wiesbaden: Springer VS.

Reichertz, Jo. 2018a. Masse, Kommunikation, Eskalation, Gewalt. In *Emotion. Eskalation. Gewalt*, ed. Jo Reichertz and Verena Keysers, 203-348. Weinheim: Juventa.

Reichertz, Jo. 2018b. Über den Imperativ der Dinge. In *Wissensrelationen*, ed. Angelika Poferl and Michaela Pfadenhauer, 108-119. Weinheim: Beltz.

Reichertz, Jo, and Carina Englert. 2021 (2013). *Einführung in die Videoanalyse*. Wiesbaden: VS Springer.

Reichertz, Jo, and Verena Keysers. 2018. *Emotion. Eskalation. Gewalt. Wie kommt es zu Gewalttätigkeiten vor, während und nach Bundesliga-Fußballspielen?* Weinheim: Beltz Juventa.

Reichertz, Jo, and Verena Keysers. 2021. Die Vereinzelungsanlage als Sicherheitsdispositiv. In *Komplexe Dynamiken globaler und lokaler*

    Entwicklungen, ed. Nicole Burzan. https://publikationen.soziologie.de/index.php/kongressband_2018/article/view/1071/1366.

Reichertz, Jo, and René Tuma, eds. 2017. *Der Kommunikative Konstruktivismus bei der Arbeit*. Weinheim: Juventa.

Soeffner, Hans-Georg. 1991. „Trajectory. Das geplante Fragment. Die Kritik der empirischen Vernunft bei Anselm Strauss. In *BIOS* 1/91: 1-12.

Strauss, Anselm. 1991. *Creating Sociological Awareness*. New Brunswick: Transaction Publishers.

Strauss, Anselm. 1993. *Continual Permutations of Action*. New York: de Gruyter

Trotha, Trutz von. 1997. *Soziologie der Gewalt*. Göttingen: Westdeutscher Verlag.

Tuma, René. 2016. *Videoprofis im Alltag*. Wiesbaden: VS.

Turner, Victor. 1995. *Vom Ritual zum Theater*. Frankfurt am Main: Fischer.

Historical Social Research

All articles published in HSR Special Issue 47 (2022) 1:
Visibilities of Violence: Microscopic Studies of Violent Events and Beyond

Historical Social Research