Video Analysis and Ethnographic Knowledge: An Empirical Study of Video Analysis Practices

Christian Meier zu Verl¹ and René Tuma²

Abstract
This paper discusses the practical foundations of ethnographically informed video analysis by investigating empirically one of the core activities of video research in sociology: the video data session. Most discussions are shaped by methodological considerations, little is known however about actual video analysis practices. By making these practices itself an object of analysis, we do show how interpretation is a social and communicative activity. In doing so, we highlight different forms of knowledge that are a resource for and topic of ethnography and video analysis. To frame our argument, we discuss the current methodological discourse on videography. Subsequently, we focus on empirical video data from video data sessions of a research network in order to discover the details of video analysis practices. We conclude this paper by highlighting our empirical findings: Video analysis is carried out communicatively by labelling knowledge, creating quotable objects through bodily reenactments, translating professional knowledge, and reassessing irritations.

Keywords
Ethnography, video analysis, knowledge, reenactment, data session

¹Department of History and Sociology, University of Konstanz, Konstanz, Germany
²Institute of Sociology, TU Berlin, Berlin, Germany

Corresponding Author:
Christian Meier zu Verl, University of Konstanz, Department of History and Sociology, Universitätsstrasse 10, Konstanz, 78464, Germany.
Email: christian.meier-zu-verl@uni-konstanz.de
Introduction

This paper discusses the practical foundation of ethnographically informed video analysis by empirically investigating one of the core activities of video research in sociology: the video data session. We are especially interested in how video analysis is conducted in vivo; thereby we focus on the bodily actions of the video analysts, forms of knowledge used to interpret video footage, and the methodological role of ethnography in video analysis. Our study is driven by the following questions: How do participants in video data sessions produce new insights into the recorded social reality? How is ethnographic and bodily knowledge used to produce these insights into social reality? How do participants in data sessions combine situated resources for interpretation?

In order to answer these questions, we perform a reflexive form of research: the activities in an empirical data session are scrutinized with videoographic methodology, recorded and turned into data to make practices of video analysis observable. This study builds on a theoretical background in ethnomethodology (Garfinkel 2002; Meyer 2019) and communicative constructivism (Knoblauch 2019). Whereas the ethnomethodological perspective highlights social phenomena as an ongoing accomplishment, the closely related approach of communicative construction extends that perspective by addressing aspects that are brought into interactions as resources, like the participants’ stocks of knowledge, institutionalized forms of communication and the materiality of objectivations. Both of these theoretical traditions are closely linked to seminal ethnographic and video analytic studies which are related to our study (c.f. e.g. Goodwin 1994; vom Lehn and Hitzler 2015). In accordance with both approaches, we understand social reality as an ongoing accomplishment or, as we call it communicative construction in terms of the applied theory, which needs to be analyzed with a focus on transient practices and actions. In this methodological sense, one’s actual conduct and its adequate sociological description are key. Thus, video and audio recordings on the one hand and practical knowledge about the investigated phenomena on the other hand render social reality accessible for sociologists.

The sociological interests in the study of research methods have a long tradition within methodological discussions, the sociology of knowledge and Science and Technology Studies. This line of research has also increasingly turned to the social sciences (c.f. Camic, Gross, and Lamont 2012; Garfinkel 2019; Maynard et al. 2002), thus reflexively focusing on their own tools and methods of knowledge production, for example, highlighting the history, effects, and politics of (quantitative) social science methods (c.f. Savage 2010). In this contribution, we promote an empirical reflection of research
practices. We show that this perspective is not only interesting for the social studies of the social sciences but also relevant for methodological debates within the social sciences.

In the next section, we focus on methodological aspects of video data. Subsequently, we give some insights into the video data session as one of the core activities in video research. In this sense, practices of video data analysis are both topic and resource within our self-reflexive empirical study of video data sessions. Thereafter, we examine four extracts from video-recorded video data sessions. We conclude our paper by reformulating our empirical findings as re-specified methodological considerations which are both self-reflexively and empirically based.

**Video Technology and Video Data in the Social Sciences**

Video data is increasingly used in a variety of research fields focusing on social interactions and related social and cultural phenomena. Prevalent video analytic methods vary in their approaches, ranging from standardized methods that rely heavily on the use of predefined coding schemes to interpretive methods. However, empirical reflection on the actual practices of video analysis are still rare.

The term video refers to a number of different data types in everyday life, as well as research usage for observational studies. They range from so called natural research recordings to elaborately produced media productions. Rather than focusing on recordings created by others, in experimental settings or media data, our focus is on research that uses recordings made by video ethnographers in order to investigate phenomena that took place in front of the camera. Because of their widespread availability, these ethnographic recordings of everyday life enable researchers to make bodily practices and communicative constructions of action visible. Our research is based on methods of sociological video analysis, that have been established internationally in recent years (c.f. Heath, Hindmarsh, and Luff 2010; Knoblauch, Schnettler, and Tuma 2014).

One core assumption of the specific version of video analysis we are focusing on—which stems from ethnomethodology and having roots in conversation analysis—is that human conduct and social reality have to be understood as a member’s practical and ongoing accomplishment. In consequence, this approach focuses on interactions, and uses video data to make these interactions accessible for research. There are specific methodological and technical particularities to video technology that we would like to point out. It allows to record mimetically audible and visually perceptible events
sequentially in order to play them back in other contexts. Video technology not only makes elusive phenomena in time and space recognizable, it also reveals different types of knowledge related to these phenomena. This knowledge is necessary to conduct and understand the practices recorded on video.

For example, Grimshaw (1982, 122) emphasizes the aspect of permanence of video data, which is associated with a particular density: it allows for a much finer and more precise observation of processes than would be possible with non-technical observation only. In contrast to other research data, such as interviews, which are used to indirectly reconstruct events, video is therefore referred to as the registration-conserving medium (c.f. Bergmann and Luckmann 1995).³

The Video Data Session as Topic and Resource⁴

Current empirical studies of knowledge and science have redirected their focus to sociology, especially to social research as a topic and resource (c.f. e.g. Antaki et al. 2008; Maynard et al. 2002). In the following, we specify the relevance of data sessions for interpretive research and briefly discuss existing studies.

Data sessions are an important focal point of interpretative work. They are part of the research lifecycle itself, starting for example, with ethnographic fieldwork, recording within the field, and finishing by producing texts about the field and fieldwork done. Even if the process is not in most cases linear and usually chaotic but also creative and context-adapting, data sessions are an essential part in most established directions of social scientific video research. Data sessions come in a variety of forms, ranging from first viewings and generating ideas with research teams to checking preliminary results with colleagues who are not yet familiar with the research in detail. The methodological literature recommends a heterogeneous composition of the group (Heath et al. 2010, 156; Knoblauch et al. 2014, 95): Usually there is at least one person who is familiar with the field of research. Often, however, it is precisely those participants of the sessions who were not involved in the field who, due to relative ignorance, are able to pose fundamental questions about the material. Depending on the method of analysis and interpretation, it may also be necessary for the interpreters to position themselves differently in relation to the material in order to introduce different knowledge stocks into the joint interpretation work.

Data sessions can be described as a form of interaction or a communicative genre⁵ in which a specific communicative problem is dealt with. The aim of the data session is both to develop initial interpretations of empirical material and to present analyses prepared by individual participants for review. Meier zu Verl (2018) reconstructs ethnographic data sessions in which
ethnographers cooperatively analyze their empirical material by highlighting, paraphrasing, and reenacting it in order to reconstruct social reality and to make its sociological relevance audible and/or visible. Ethnographers transform their material into data practically, which means that they gradually treat it as meaningful and sociologically relevant in data sessions. Therefore, Meier zu Verl concludes his ethnomethodological “ethnography of ethnography” by characterizing ethnographic research processes as observable data careers: from initial observations to empirical material to data candidates and finally to sociological data (c.f. Schindler 2018; Sormani 2020). Video analyses of video analysis data sessions are still rare. Nevertheless, Tutt and Hindmarsh (2011) deal with video data sessions empirically. Using examples, their analysis breaks down data sessions in terms of their sequential organization by the social researchers taking part in video data sessions. For the authors, analyzing and reconstructing the video means foremost a process of bodily showing and reciprocal naming of relevant elements. One important empirical finding of their study is the reenactment. Reenactments are (methodological) practices used to render video recorded practices visible again and also to make embodied knowledge practically noticeable for oneself and observable for others at the same time. As the authors show, analytic activities are not only a linguistic or purely cognitive assignment, but it is rather the bodily showing and reenacting that plays a central interpretative role in video data sessions. While they describe in detail how social researchers conduct practical reenactments to analyze and interpret their video data, empirically not much is known about the actual interplay of video analysis practices (such as reenactments) and embodied knowledge in video data sessions. Here the communicative constructivist perspective contributes by explaining the relevance of such methods by paying particular attention to the specific problems that the participating subjects are confronted with: Knoblauch and Schnettler (2012, 353) emphasize that “[t]he analysts do not only mirror an activity identified in the fragment. They translate the indexicality of the situation watched on the video into their ‘situative indexicality’.” Thus, reenactments are bodily practices to understand social reality not only through cognition but also through one’s own (inter)corporeal experience in the first place. Extending the domain of academic scientific analysis and also comparing different cases, Tuma (2017) studies the communicative interpretation forms of “vernacular video analysis” (e.g., in market research, police work, or sports analysis) and shows how they are embedded in a “work arc” that is specific to professional fields.

In this paper we take up those findings, but we add upon this research as well, by systematically scrutinizing a case for the different forms of knowledge that are invoked in video data sessions.
We consider reenactments and other communicative forms exhibited as resources and solutions to specific practical and communicative problems of joint interpretation.

**Practices of Video Analysis and Different Forms of Knowledge**

Based on our activities as founding members of the German researchers’ network videoanalyse.net, we recorded some data sessions which were conducted from 2015 to 2019 and created a wide-ranging corpus of data sessions. These recorded data sessions vary in disciplinary orientation of the participants, who have slightly different theoretical backgrounds (from symbolic interactionism, communicative constructivism, and ethnomethodology to sociolinguistic), methodological approaches (video-ethnography and multimodal video analysis), and research topics. It is an essential characteristic of such meetings that researchers with different perspectives and research interests meet and discuss fragments of data. Nevertheless, most participants agree on some basic principles that are common across the field of interpretive video-analysis. For this paper we focus on the reflexive analysis of one specific video data session. We recorded the approximately 90-minute data session in which a short video fragment was presented and replayed. The data session was video recorded by us and therefore reflexively turned into a further set of video data that we then subjected to a meta or reflexive analysis, which forms the basis of this paper. The following extracts from one recorded data session have been selected as they highlight aspects that are important to our research question and serve as examples that allow us to display our findings in a comprehensive way. An important self-reflexive aspect of our analysis is that we actively participated in the analyzed video data session ourselves. This might be considered a methodological problem in some traditions of social research; however, since our argument will be empirically elaborated in the following, our participation in the investigated analytical context provides us with essential (ethnographic) background knowledge, to understand video data sessions not only in general but also to understand this video data session in particular and render it understandable for other colleagues. Methodological considerations about being or going native as sociological video analysts might be seen to cause friction. This was reflected in our (a) specific analytical perspective of second-order estrangement, that is, compartmentalization via meticulous fine-grained sequential scrutiny of our own comportment; (b) temporal distance between situations (previous
video data session and our current video analysis); and (c) analysis and communicative validation in other (meta-)data sessions. Therefore, we systematically apply the videographic methodological demand for ethnographic participation within our own field.

*Description of the Ethnographic Video Footage*

The recorded video data session was part of a two-day workshop. One of the participants (Eric’s all names pseudonymized) had prepared some video data that was part of his current research (meanwhile published in Albrecht 2016) and that is discussed within the group. Most of the participants were neither familiar with the data nor the project itself; however some (e.g. Andy, see further) had been discussing the project with Eric beforehand, including other fragments of data pertaining to it. Eric had been working on his personal videographic study (ethnography using video), which focused on the measurement and evaluation of creativity in the context of engineering work, studied in academic teaching and examination. Thus, he is the only one who is well acquainted with ethnographic knowledge in the field of engineering. The fragment he presents is part of a bigger data corpus that consists of observations, fieldnotes, and video recordings of a specific project-oriented teaching format. The specific video presented by Eric focuses on a situation where a small group of students are discussing the results of an engineering project with the responsible teaching assistant and student tutor who supervised their project. The students are using their notebook and a CAD software, which is typically used for visualization in engineering (i.e. technical blueprints). This presentation is important, as the assistant and tutor are using it as the basis for grading the project, evaluating different aspects of the work, including the creativity of the solution being explained by the students in this situation. Eric was especially interested in the process of evaluation, and how teachers also understand, evaluate, and assess creativity based on situated communication. Figure 1 shows a transcript of his video footage that Eric has presented to the group.

This transcript and the accompanying video recording were analyzed by the participants in the data session. In our next four analyses we will show how Eric’s co-participants study this particular fragment of interaction without having the same knowledge of engineering as the ethnographer Eric. We will reveal how these participants achieve new knowledge about this video recorded interaction, which is situated in their analysis as a shared, ongoing, and bodily accomplishment.
Common-sense Knowledge, the Body and Visibility

The following reflexive analysis of the data session shows how knowledge is practically located and marked as missing knowledge by reenacting gestures, referencing, and highlighting passages of transcript, and making one’s own limited knowledge about professional activities public.

After an introduction to Eric’s PhD project, his video footage is played several times. Eric asks all participants of the video analysis group what might be of sociological interest about the video in general and more specifically how examiners practically assess specific achievements like the creativity of their students. The following extract displays an early walkthrough of the video footage and its transcript. All participants are sociologists, and no one comes from either an engineering or a natural and applied science background. Eric as an ethnographer in the field of engineering therefore has unique explicit and tacit knowledge about this particular
video recorded interaction and also about teaching and examining engineers in general.

From line one to line four, Carl begins his reconstruction of the recorded interaction by formulating individual observations that were acknowledged by Eric uttering continuers (lines 2 and 4). These observations were highlighted by Carl as personally interesting places within Eric’s video data (like “so what is interesting for me” line 1). While doing so, Carl orientates himself to the written transcript (line 3), the video recording (Figure 2), and to the video recorded participants, their knowledge, their relevancies, and their interactional orientation to certain objects (line 5). His orientation to the video recording can likewise be seen as a polysemous way of reenactment:
Carl is directing his colleagues’ glances onto the display for analytical purposes. In addition, he uses the same gestures (pointing gestures) used by the research participants in Eric’s video footage to direct glances. Thereby, the practical purpose of pointing and directing others’ gazes is practically experienced within the data session and at the same time discussed sociologically (lines 6 and 7). However, this type of reenactment was not explicitly addressed by Carl or anyone else, probably because it could not directly be linked to the concept of creativity.

Afterwards, Carl questions his personal knowledge to fully understand the video recorded actions and finishes his speech by pointing out that there is no assessment formulated (line 8). With the following upshot, Carl reinforces his observation that there is no assessment formulated by giving his colleagues empirical evidence. This form of evidence is achieved by reenacting and verbally highlighting one phrase that is accessible to all through the transcript. However, Carl does not come up with analytical terms. In addition, Carl and Eric were subsequently clarifying which student was actually examined (not part of our transcript, line 10). After this thematic insertion, Andy claims to have more knowledge about the shown video because he had already seen more of Eric’s data, and states that there is an assessment which is actually taking place but is without any readable evidence in the written transcript (line 11). In this sense, certain familiarities with (watching and showing) video data are analytically important as well. Eric confirms Andy’s statement afterward (line 12). Furthermore, Carl confirms Andy’s observation, which partly reformulates the observations he himself had previously addressed. On the other hand, his formulation also questions Andy’s statement (that there is an assessment at all) by asking Andy directly: “and where can you see it?” (line 13). Andy locates his assumption that there is an assessment in line eight in the written transcript and nearly simultaneously begins to play back the video (lines 14–16).

We are reading this transcribed interaction as an ongoing accomplishment of one specific video analytic group. Their analysis is driven by a particular sociological interest in creativity assessments, that is sensitizing the specific form of analysis. Carl describes what he considers to be seen in the video recording by highlighting some moments of the recorded interaction in order to understand their sequential emergence. In this way, he addresses the question of how creativity can be assessed co-operatively without, however, formulating a less descriptive and more conceptual answer for his co-participants. Here, Andy applies a further analytical
strategy. He does not transform the observed interaction into his own words, instead he draws on and combines different resources of interpretation, that is, the transcript and the recording itself by playing the video. In doing so, he claims a different epistemic approach. This extract refers to a number of sense making practices arranged in a specific order: (a) The highlighting of possible sociological relevant places in the written transcript, (b) blurring the line between the video recorded situation and the current situation by reenacting gestures with similar interactional purposes, (c) linking the reenacted gestures back to the verbal by phrasing it in conversation analytic terms, (d) calling attention to a gap between these previous empirical observations and the thematic framework previously introduced by Eric, and (e) disagreement within the video analytic group and therefore going back to the video by replaying it once again to locate a possible position where the assessment takes place. Knowledge, and especially missing knowledge, is addressed as a key interpretative feature by the observed members of the video analytical group in order to formulate sociologically relevant interpretations. It is not framed as a problem of specific ethnographic knowledge that is missing in the group. The problem emerges by failing to use one’s own knowledge meaningfully; it also arises by questioning and naming the missing knowledge. Methodologically speaking, it is not only that the understanding but also (as we have shown) the non-understanding is an ongoing accomplishment, which is ordered sequentially and achieved practically.

**Rendering (Bodily) Knowledge Recognizable**

With the next analysis of video analytic practices, we show that professional knowledge can be rendered accountable by reenacting gestures and simultaneously commenting on them. Both practical and communicative interwoven activities enable the ethnographer Eric and his colleagues to (re)activate bodily knowledge and to make sense communicatively. This sensemaking activity in terms of reenacting parts of video footages and commenting on the current reenactment is related to other video analytic practices such as watching and commenting with a certain difference: the bodily reenactment is a preliminary stage before describing something sociologically.

After playing the video once again, the group members begin another attempt to interpret how the video recorded participants (tutors) perform their assessments in practice.
Extract 2. WS-DS1 (13:48–14:47 min.).

Andy: yeah so my assumption is that he- that- that just uhm- he reconstructs what he meant here he had understood this construction uhm uh anyway it is screwed over there and he reconstructs that by his hand gesture and uhm and for him it is something which he hadn’t expected, right?

Andy and Carl

And uh ah I uh we have a solution and not a standard solution therefore he is asking again because he somehow before probably- I assume that he has stuck it on or what did I know uhm that he says then oh, I see

Carl: [( )

Andy: mmm

Carl: there is no oh, I see or something like that this would mean that there is a change of state token okay

Video: ((play-back))

Andy: yes

Carl: but he doesn’t say oh, I see he- he somehow makes descriptive how it is done in this picture- this is embodied by him again so I think it is- it’s not much more-

Andy: mmm yes

Eric: so it is indeed not really readable in the transcript- only if you have the footage as well can you somehow determine that there is an assessment happening

Shortly after this video playback, Andy formulates an interpretation, which is ordered as follows: (a) He indicates that his interpretation is part of his own intuition (“yeah so my assumption is”; line 1), (b) he formulates an overall interpretation of all the video footage (following line 1), (c) at the end of this first statement, he asks his co-participants for their agreement (“right?”; end of line 1), and (d) he continues at the beginning of his next turn by assessing (like a tutor) the technical achievement of the students (“and uh ah I uh we have a solution and not a standard solution”; line 2), and while doing so reenacts a specific gesture (Figure 3). In this way, he indicates verbally and gesturally a change of perspective: from sociologist to participant (skilled engineer). This reenactment by Andy not only creates a quotable object but is also a reenactment of different “roles”: While Andy is reenacting a specific gesture of one of the tutors, his co-interpreters become observers of the (reenacted) gesture itself once again. Following up this assessment, Andy (as a sociologist but implicitly framed as a skilled engineer) relocates his interpretation of an
assessment within the video recorded interaction (following line 2), thereby empirically justifying his interpretation. In order to do this, he also verbally reenacts the supposed assessment by changing his voice (“he says then oh, I see”; line 2). Carl disagrees with Andy by pointing out that the enacted tutor’s voice cannot be located in the video footage (“there is no oh, I see or something like that this would mean that there is a change of state token, okay”; line 6). While Andy tries to locate something in the video, he plays it again briefly (line 7) and confirms Carl’s objection afterwards (line 8). Once again, Carl points out that there is no “oh, I see” (line 9) and offers his own interpretation instead. This is questioned by Andy (line 10), followed by Eric’s statement reinforcing that there is an assessment happening that “can somehow” be discovered (line 11). Eric does not reflexively label missing knowledge but rather labels it implicitly and as part of a video analytical discovery.

Most importantly, this extract shows that knowledge (this applies to common-sense knowledge as well, see earlier) has to be (bodily) formulated in this situation to develop analytical and communicative outcomes in data sessions. Andy renders his various interpretations recognizable for himself and his co-interpreters by presenting analytical frameworks (“standard vs. creative solution”), giving empirical evidence (“anyway it is screwed over there, and he reconstructs that by his hand gesture”), and finally by reenacting embodied practices and communicative actions (fig. 2 and line 2 “he says then oh, I see”). These practices and actions make his interpretations livelier and more experienceable for everyone in the room. It also activates the participants subjective knowledge, for example about making an assessment—even if Andy’s interpretation and reenactments are questioned by the others. This extract shows that knowledge and its various forms are mediated not only through the video footage and the transcript itself but also through communicative and embodied practices that transform practical knowledge into recognizable, visible, observable, and in the end accountable forms of certain members’ actions (e.g. actions of engineering tutors). These practices are methodologically underestimated but there are—as we have shown—practically important to gain first-hand insights into social reality. Reenactments are those first-hand devices that can help to produce sociological knowledge on the basis of one’s own performed embodied knowledge.

**Technical Language, Skills and Ethnographic Knowledge**

The following analysis shows how apparently ordinary language can be transformed into specific technical language through a particular variation in perception. This variation is practically accomplished by the ethnographer Eric, not only through bodily reenactments but also by referring to (common) semiotic resources which seem to be more available to his co-present colleagues.

01 Dana  What I notice again here is that there’s almost no expert lingo. Is it really like that, that they generally communicate at this almost (colloquial) [-] level? For instance this “I checked, there’s plenty of room”. That sounds at first like they don’t really what that, there, is, I’d say OK, they don’t know their way around it that well, is what I would say.

02 <GROUP LAUGHS>

03 Dana  Hmm I’d say that they can’t express it any differently than to say there’s enough room, wrapping it in everyday speech. So, I don’t want to judge but the question do they talk all the time like that or is that just a very special sequence (--) where they’re trying to somehow break it down, or what?

04 Eric Most of the time they talk like that

05 Dana  hm

06 Eric  But you still don’t understand when you don’t know what it’s about, but it is (.) I would say strongly indexical, i.e. they’re talking about mounting and screwing on, that refers to very specific actions. Mounting is decidedly not screwing on. So, anyone with an everyday understanding would say that mounting is something that includes screwing on, but it’s not that, mounting, that means somehow a complex connection, which for the machinists here it’s already about which tools they need to do it.

If a special tool is needed, when these materials are needed, so perhaps it’s a screw that one can only screw in once and then it breaks when one tries to take it out again, so this mounting is a very complex problem construction and there its about this screwing on <TURNING GESTURE> And then simply this screwing on of this little screw still <UP AND DOWN MOVEMENT WITH FINGER THEN COMBINATION WITH SCREWING MOVEMENT> put on from the other side, which has also to do with counteracting, thus the screwing on has to be fixed so that it doesn’t unscrew itself in such a rotating system (.) and ehm this little screw is only there for that.

Later in the data session, Dana addresses the specific language used by the students and also the tutor in the recording. From her point of view, she classifies their way of speaking as a mundane form of language and voices her confusion about the lack of professional terms that she would have expected in this context (“What I notice again here is that there’s almost no expert lingo. Is it really like that, that they generally communicate at this almost [colloquial] [-] level?”; line 1). She roughly quotes with a short
pre-sequence a sentence “For instance this ‘I checked, there’s plenty of

room’.” (line 1) from the data played before, which is also represented in

the transcript (see Figure 1, line 10) that is available to her, as well as to the

other members of the interpretation group. Her quoting, as well as the pre-

vious reenacting by Andy (Figure 3) can be seen as an epistemic practice to

produce audible and visible evidence for her co-participants and to use this

situated produced evidence as proof for previously introduced interpreta-

tions. Her remark (line 1) that mundane language can be understood as a

missing expertise of the engineers evokes a short burst of laughter among

the group. She asks, addressing indirectly Eric, who, as mentioned before,

knows the whole data corpus and was responsible for ethnographic research,

whether this is specific for this short sequence or a general characteristic of

the situation. Eric begins with a quick confirmation implicitly based on his

ethnographic knowledge that this use of language can be generalized

(“Most of the time they talk like that”; line 4). He then goes on to explain

that the used terms are not to be seen as vernacular; rather they are present

forms of professional language that can only be read and understood with

specialist knowledge (by mentioning, i.e., the concept of indexicality “But

you still don’t understand when you don’t know what it’s about, but it is. . ..

I would say strongly indexical, i.e. they’re talking about mounting and

screwing on, that refers to very specific actions.”; line 6). He elaborates on

this by explaining the differences between “montieren / mounting” and

“aufschrauben / screwing” (line 6).

Particularly interesting in Eric’s prolonged explanation is how he relates

the specificity of the seemingly vernacular language to the specific context of

construction work. Using the example of the fixation of rotating machine

parts, he addresses the complexities of the engineering process. His explana-

tion draws not only on his conceptional knowledge of engineering, which he

articulates in the description of certain examples of problems and the details

that engineers need to consider on the basis of their instruction (blueprints

and the use of specific tools), but he also makes certain “instructed action”

(Garfinkel 2002, 197–218) accountable by using a specific hand gesture (fig.

4, 5) to accompany his speech in order to make his expert knowledge visible

(c.f. Sormani 2014, 123–4). By enacting “mounting” and “screwing” with

turning movements of his hand with open fingers, he makes the specific engi-

neering knowledge not only accessible through language but also available in

a performative imaginative space, bodily, by enacting the material and three-

dimensional structure and working of the machinery. He enacts the main

screwing with his left hand and combines the gesture with a specific finger

movement of his right hand, thus performing the fixation of smaller screws

that fix the entire connection.
First and foremost, the sequence has shown how field knowledge (i.e. knowledge of engineering) is brought into video data sessions as a specific use of natural language and its indexicality that must be specifically understood within this field. To achieve this, the explanation of Eric the ethnographer is necessary: he explains the utterances in the transcript based on the specific acquaintance with the engineers’ expert knowledge, which he has acquired during his research. Second, he uses a specific form to make his explanations accountable through the field-specific distinction between instruction and instructed action by explaining some serious problems using blueprints that illustrate how to assemble machines. Third, the practical examples are not only explained verbally but also performed bodily by reenacting the assembly. In the case of this data session, the role of the bodily knowledge that is enacted is part of the usual reenactment of movements on screen. The group focuses on this dimension of interaction explicitly, due to their theoretical and methodological orientation and also because the research question that Eric presented in the beginning was aiming at this dimension. In other words, enactments can be used as devices of practical reasoning or explaining in an attempt to make ethnographic knowledge accountable, like reenactments that are not only practical devices of epistemic fact finding and plausibility checks depending on its sequential organization (as we have shown earlier). If a reenactment occurs at the beginning of a statement and is followed by a (pre)conceptual description of the reenacted phenomenon, it is used epistemically to produce new findings. If the sequential order is reversed, the reenactment is used to check the plausibility of the preceding conceptual description of the identified phenomenon.

**Making a Perception Visible as a Fallacy**

We have already shown that the analytical knowledge that is necessary to interpret video footage is continuously produced, assessed, relativized, etc. among all participants to identify and describe phenomena of possible relevance. With our last analysis we showed that situationally produced findings can be revised by co-participants much like any other resource for video analysis.

At the end of the data session, some participants discuss more focused conceptual implications of the specific gesture displayed in the video. This gesture, that Eric enacted in his practical reasoning (Figures 4 and 5), becomes the explicit subject once again, albeit now directly associated with the observed movements on screen as an iconic gesture of screwing. This short sequence of hand movements is repeatedly played by Andy and the group focuses on it.
In this last extract from video, Andy states why he does not understand the actions of the participants on screen by making a description that unfolds close to the temporality of the video footage and helps him to make
his non-understanding comprehensible for his co-participants (line 1). At first Andy introduces his following description as part of a personal confusion, which enables him to go back to a previous point of analysis (“What still bothers me is actual”; line 1). For his verbal description, Andy also uses the playback of the video as an evidencing instance and does not reenact any gestures from the material (as we have seen earlier). In this case Andy directs the attention of his co-participants with a pointing gesture to parts of the video that are relevant for his description (Figure 6). Then, at the end of his turn, Andy announces his expected course of the interaction, which does not correspond to the recorded course (“I would have actually expected that he would sit down and somehow turn the thing around in the computer”; line 1). He tries to understand the actions of the participants in context and uses his own pragmatic expectations of the course of events (as established in the former sequence 4.3) to hermeneutically project other possible solutions to the participants’ problems of action. After this announcement of expected events by Andy, Bert begins his statement by agreeing on Andy’s expected assumptions and then gives a detailed verbal description of which hand does what, by whom (“That’s what he’s doing with his hand, isn’t it? (...) Like, the left hand that he’s rotating, that’s not the one from the guy in white, I mean, in the white t-shirt, rather from the guy at the computer”; line 2). While Andy’s description still implicitly assumes one gesturing person, with Bert’s description it becomes public and explicit that the two gesturing hands are from two different persons. In other words, in the video you can see two gestures that are synchronous but not directly related to each other. It is somehow a visual illusion that is caused by perspective of the filming camera. After Bert’s adjustment, Eric reenacts the same two gestures with his own two hands and begins to confirm also verbally this (new) observation (line 3). Directly following Eric’s turn, Andy articulates a “change of state token” (Heritage 1984) and thus makes public his change of stance to his co-participants (line 4). Following this, Bert again describes the two gestures and goes into more detail about the belonging of the two individual hands (line 5). Eric also reenacts the two gestures, making both visible again, while highlighting the misreading of the video (Figures 7 and 8). At the same time, Carl makes the two gestures (Figure 8), while Andy still looks at the screen. Carl also confirms this new situated understanding by saying “that is the other hand” (line 7) and then reenacting just one of the two gestures in the reverse direction (Figure 9). Following Carl’s confirmation, Andy displays his amusement over his misunderstanding by performing an overly expressive “smack to the head” gesture accompanied by a wide grin (Figure 10).
This extract not only shows epistemic practices of reenacting and of replaying specific movements that are on the screen but also that they can be analytically and sequentially put into an order by participants as an ongoing achievement to discover a visual fallacy. As we have seen, this practice of discovering a fallacy follows a similar sequential organization of turns as seen in the previously analyzed extracts. But this time it is much more distributed and carried out in a certain order by four co-participants. For example, Andy describes the interaction he observed in the video and tries to make his description plausible by replaying the video. Or Bert also describes this interaction, followed by reenactments of Eric and Carl. These descriptions, but also these reenactments are produced context-sensitive and therefore follow different relevancies. More importantly, they deal with various practical and epistemic problems that emerge in the data session itself. Insights are therefore, so to speak, produced from within. The outside interactions, recorded by the video, are only being made recognizable, analyzable, etc. through the practices of the participants in the here and now of their data session. Even if the production of knowledge in data sessions is based on fallacies, there is no corrective from the outside, but only those that are practically produced from within. This also holds true for the brought along ethnographic knowledge, which must be made accountable in order to be become relevant for the situated production of social knowledge.

Conclusion

We conclude our video analysis of video analytic practices by emphasizing our findings and addressing different aspects of common-sense, professional and ethnographic knowledge as a situated achievement of joint video data sessions. We want to emphasize the sequential organization of our analyzed interactions, with which the observed participants developed and proved their new insights practically and co-operatively. The video data session is discussed as a central part of research in the methodological literature (c.f. subsection 2), but without including empirical descriptions of the practice to develop new methodologies. Our empirical and self-reflexive video analysis of video analysis helps to uncover various practices of doing video analysis. It is an ongoing process and it brings together methodology and empirical findings in a new way: an empirically based video methodology. Our main argument is that practices of data analyses are observably interwoven with different forms of knowledge, which are systematically integrated and checked against each other. This is an aspect that is highlighted by philosophical traditions like phenomenology and hermeneutic but also by ethnomethodology and communicative constructivism. They
address sociality and knowledge even more explicitly in contrast to other less ethnographic traditions of social research.\textsuperscript{11}

Our analysis of the practice of video analysis shows that common-sense and ethnographic knowledge instruct our perception of seeing (and hearing) video data and can therefore to be understood as enabling resources. This is not a purely theoretical problem; rather, it is a practical problem that participants in data sessions have to deal with by employing practical strategies, such as labeling assumed missing knowledge, requesting and making explicit more information about the ethnographic field that is being examined (extract 1). This pertains to not only clear discursive questions but also more subtle aspects such as tacit or bodily forms of knowledge that are essential to make sense of video recordings. Embodied knowledge as visible practices is highlighted practically by reenacting it. It is of methodological importance that these reenactments create quotable bodily objects and make bodily experiences analytically available. This availability is crucial for understanding the organization of a data session, because we can observe temporarily different roles within the video analytic group in the course of their interactions: (a) the reenactor who is bodily performing the visible movements and (b) his colleagues who observe his or her transformation as specific observers. Reenactments analytically transform not only the interpreter who reenacts them but also his or her co-interpreters as observing witnesses of actual reenactments (extract 2). They are methodological and practical important devices—not only for highlighting bodily knowledge but also for tentatively producing findings that are initially expressed bodily and then have to be transformed into words subsequently. Furthermore, they can be used as device of proof for previous descriptions and interpretations through practical validation (extract 3).

The video data is interpreted and gradually reinterpreted, by referring to the ethnographic knowledge gathered during fieldwork, which is then used for a context-specific interpretation of (verbal) articulations and turns. The meaning of the (spoken) statements in focus is interpreted by the group on the basis of the field-specific insider’s knowledge, for example, the specific meaning of language and terms (extract 3) that appear to be mundane, but which are reinterpreted as very technical and as existing within one specific profession. However, verbal expressions are not disentangled but rather are understood in relation with bodily reenactments that have previously been established as quotable.

The participants in the video data session apply a pragmatic motive based on the reconstruction of situational relevancies and the available resources. The sequential order of the ongoing action is reconstructed based on the alignment of the actor’s perspective. Irritations and unfitting turns, combined with repeated viewing of the sequences, can lead to a reassessment of
formerly established new findings. This is shown in extract four, where, through the interpretation, it is realized that there is a visual fallacy being reproduced by the reenacted gesture; the sequence is then reinterpreted respectively.

The interpretation is guided not only by the reconstruction of the actions but also by the research questions at hand,\textsuperscript{12} that is, by the sociological relevance and therefore by a specific professional vision of professional sociologists (extract 1 and footnote 6). Even if this aims at the reconstruction of members’ knowledge, this interest is connected to sociological thinking.

We want to emphasize that our empirical findings methodologically reinforce the relevance of ethnography, ethnographically acquired knowledge, and common-sense knowledge for the practical conduct of video analysis. If we take Schütz (1953) and Garfinkel (2002) and their postulates of adequate sociological descriptions seriously, we have to take actual video analytic practices into account in order to reformulate our methodological considerations. These reformulated methodologies can be achieved—as we have shown—by expanding them empirically and by describing reflexively methodological practices such as \textit{topic} and \textit{resource}.

The participants in the video data session combine different forms of knowledge by confronting and examining the data with each other. New ideas are brought up on the basis of sociological questions (extract 1), and reenactments, as quotable epistemic objects, are shown again on screen and made accessible through performance and visibility (extract 2 and 3). They are put into question by the researcher’s ethnographic knowledge. If these different forms of knowledge are brought into unison, it is likely that the participants are either accepting the situated interpretation as sound or revising it (extract 4). Even if interpretive video research is not based on a general logic of testing, a logic of falsification can be employed in a complex manner during video data sessions.

\textbf{Authors’ Note}

This manuscript is a contribution to a special issue of JCE entitled Practices of Ethnographic Research, guest edited by Andrea Ploder and Julian Hamann.

\textbf{Acknowledgments}

We thank Sandrine Gukelberger, Laura Keesman, Hubert Knoblauch, Christian Meyer, Phie van Rompu, Theresa Vollmer, Don Weenink, and both guest editors of this special issue for the insightful comments on the drafts of this paper. Furthermore, we thank Elisabeth Schmidt and Jacob Watson for the meticulous proofreading and Marie Tuma the fine drawings. We are appreciative to all participants of our research network videoanalyse.net for their participation and to the anonymous reviewers for their detailed suggestions. However, remaining mistakes are ours.
Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Christian Meier zu Verl https://orcid.org/0000-0002-4611-5984
René Tuma https://orcid.org/0000-0001-9672-1694

Notes

1. This research can be considered (self-)reflexive as the authors of this paper are part of the recordings and have been participating in the social phenomena that is being studied. This peculiarity however, need not be seen as problematic for this research and the methodological approach taken. It can rather be understood as a resource for interpretation, as both methodological considerations as well as our empirical findings show that field-specific participants’ knowledge is integral for context-specific and meaningful interpretations of the situations we are going to portray in the empirical part of this paper.

2. We are aware of the slight differences and standpoints in the discussion between theoretical traditions; however, the cooperation between us as researchers with slightly different angles and terms (e.g. preference of the term practice in ethnomethodology versus the term communicative action in communicative construction) leads to productive insights. In this paper, this leads to some formulations that cover both approaches by juxtaposing both terms.

3. There are different notions of data within interpretive research. Ethnomethodology has the notion that data is an “exquisitely detailed reminder(s) of what we already do, as a matter of course” or “a source of insight into what we ‘already know’” (Lynch 2002, p. 535); it is (or “they are” depending on if you used plural before) also a hermeneutic and phenomenological aim at methodologically explicating the specific subjective knowledge tied to any kind of interpretation of data (c.f. Knoblauch & Schnettler 2012).

4. We refer directly to Zimmerman and Pollner (1970) with the distinction between topic and resource.

5. A concept used to highlight that reoccurring forms of communication can be considered as a kind of institution, that has been established within a community and allows to deal with communicative problems (c.f. Bergmann & Luckmann 1995).

6. The differences between data sessions conducted by groups with a more homogeneous theoretical background and meetings with diverse theoretical
and disciplinary orientation would be an interesting research topic for future research. Related to the topic of this paper, especially questions of how different traditions of research deal in practice with ethnographic or knowledge in interpretation differs to a large degree. This has an effect on the specific relevancies that are displayed in analysis, that is, the questions of interest to the participants when analyzing specific aspects of the data. We do not focus on this aspect, but it sparks interest for further research.

7. We selected this video data session for two reasons: (a) The ethnographer presenting his video data is at the time of this data session in the middle of his PhD thesis and is therefore no longer a beginner, and (b) we are interested in taken for granted practices of video analysis which emerge in many video data sessions due to their commonplace structure.

8. Together with other colleagues, we analyzed the video footage about one particular video data session that we will present subsequently. The meta data sessions were again video recorded by us and turned into data itself. These “meta-meta” video data analyses are not part of this paper but could be the possible basis of another level of empirical self-reflexivity.

9. All transcripts have been translated from German into English.


11. This highlights the problems that are present when using “found video data” rather than self-produced videos (c.f. Nassauer & Legewie, 2018).

12. The observations and interests of other participants often lead away the initially formulated question(s) of the presenting researcher and are sometimes only loosely connected.

References


**Author Biographies**

Christian Meier zu Verl, is a research associate at the Department of History and Sociology at the University of Konstanz, Germany. He is the deputy coordinator of the network *Dis-/Abilities and Digital Media*, funded by the German Research Foundation (DFG). His recent publications include *Daten-Karrieren und epistemische Materialität* (2018), *Die alternde Migrationsgesellschaft* (2020), and the special issue *Ethnomethodology and Ethnography* (Ethnographic Studies, 2020).

René Tuma, is a research associate at the Department of Sociology at Technische Universität Berlin, Germany. Currently he is a postdoc researcher in the ERC Project Group Violence at the University of Amsterdam, Netherlands. His recent publications include *Videoprofis im Alltag* (2017) and *Videography: Introduction to Interpretive Videoanalysis of Social Situations* (2014, with Hubert Knoblauch and Bernt Schnettler).