

René Tuma
Berlin Institute of Technology, Germany

The (Re)Construction of Human Conduct: “Vernacular Video Analysis”

Abstract Video technology became available in the 1960s and massively diffused into nearly all institutional spheres during the following decades (Zielinski 1986). It is not only used for provision of movies and entertainment (Greenberg 2008), documenting, and recording of events in the semi-professional and private domain (Raab 2008), but also for the analysis of human conduct in psychology and education (Mittenecker 1987) and in sociology (Knoblauch et al. 2006; Knoblauch et al. 2008; Kissmann 2009; Heath, Hindmarsh and Luff 2010). In this paper, I am going to argue that with the availability of video technology a form of discursive practice of interpretive (re)construction of knowledge has been established in a variety of vernacular fields of practice. Based on the available literature and an empirical example taken from a public demonstration of such a vernacular video analysis, I will show some elements of communicative practice that allow for detailed analysis of visual knowledge. Furthermore, I will discuss methodological issues, as to be approached, and which elements can be found that constitute the communicative (re)constructive processes of analysis.

Keywords Vernacular Video Analysis; Ethnomethodology; Interpretation; Interaction; Video; Workplace Studies; Visual Knowledge; Reflective Methods; Scientific Practice; Experts

René Tuma, M.Sc., is a researcher and lecturer at the Technische Universität Berlin and currently working on his PhD Thesis in the field of video-analysis.

email address: rene.tuma@tu-berlin.de

Background

From the beginning of the 19th century, when photography and film had been established, researchers have started to capture movement and make ephemeral phenomena of human conduct accessible to the human eye. One of the early famous examples are the images tak-

en by Eadweard Muybridge (Brookman 2010), who experimented with the fixation of movement of animals and human beings into photographs. The images have been used to analyze and optimize human movement in different spheres: very early on, Ford used recordings to rationalize the work processes in his factory in order to optimize the organization of labor. The psychologist, Kurt Lewin, used film recordings in 1923 for the analysis of behavior in conflict situations; his student Gesell also published a book about film-analysis as a methodology for the scrutiny of human behavior (Thiel 2003). Film has also been used in anthropology (Ruby 2000). Very famous examples are the studies of the Balinese Character by Bateson and Mead (1942). In the study of human micro-movement (kinemics) Ray Birdwhistell (1952) pioneered with his microscopic approach, scrutinizing the movements of a smoker, while the Paolo Alto Group studied interaction in interviews (Bateson 1958). One might also recall Ekman and Friesen (their form is the basis for a popular TV-series called “Lie to me”), who developed a method for the fine grained analysis of facial expression (2003). Histories of the use of film and video in research within the social sciences remain (Heath et al. 2010; Knoblauch, Schnettler and Tuma 2010; Erickson 2011) rather short and are usually part of methodology introductions, but highlight the growing importance of audio-visual recordings.

The historical development of the connection between film and science has been studied by Reichert (2007), who describes it as a dispositive that brings together those two spheres and

– by generating this specific dispositive – forms our knowledge (especially in sciences of the human). He argues that, at least from the 1960s onwards, the apparatus of the cinema has been understood as a tool for the transformation and organization of psychological dispositions and structures of the gaze. Referring to a number of examples (e.g., film in anthropology), he shows how especially film, but later also video is used for observation, recording, demonstration, instruction, and optimization of human activities. The author elaborated on the thesis that cinematographic practice does not only affect the production of knowledge, but constitutes it. In his book he discusses a number of deployments of film as medium of scientific research and presentation. He also has one example where it is applied for analysis: the Stanford prison experiment. However, due to his historic approach, he has no access to the interactional practice in which video is used, but rather the product. Reichert extracts from the visual products and the context, what is made visible and which forms of power-relations are contained in a specific form of technologically crystallized knowledge. For the understanding of the practice of those new technologies historical studies are very informative, but as sociologists we now should look at the interactional situations, in which actors do put the recordings to visual practice.

Assuming, that the availability of those recording technologies can change the way of perceiving the world and the social environment in a non-technologically determined way, one has to ask how and why specific images/

recordings and by that understandings of human action are produced as communication. Following the questions raised in the program for the sociology of visual knowledge (Schnettler 2007) and in the Science and Technology Studies (STS) field focusing on visibility (Burri and Dumit 2008), I argue that the performative forms in which visual knowledge is actively produced, engaged and distributed should be the focus of our scrutiny. Revising the literature available in the field, one clearly finds a large number of different studies about images and visualizations in science (for an overview see Burri and Dumit 2008), but there are only a few studies on video technology.

Professional Vision

One of the few studies of video in use (or as I call it: “vernacular video analysis”) has been presented by Charles Goodwin in his papers concerning “Professional Vision” (Goodwin 1994; 2000). The author deals with the well-known Rodney King Trial at a Californian court. In the 1990s, this case of police violence that had been filmed by a bystander was the starting point for conflict and public outrage.

In his articles about professional vision, Goodwin identifies three general practices as essential components for the production of visual knowledge:

- “Coding Scheme:” Transforms the world into the categories and events that are relevant to the work of the profession;

- “Highlighting:” Dividing a domain of scrutiny into a figure and a ground so that events relevant to the activity of the moment stand out;
- “Graphical Representations:” External representations of distinctive characteristics of the material world to organize phenomena.

Reconstructing the Rodney King trial, Goodwin interprets how a member of the police (Police Officer Sgt. Duke) delivers his interpretation of the recorded beating in front of a jury. He describes and interprets the movements of Rodney King, laying on the ground as aggressive – to be more exact – as starting to be an aggressive movement, which legitimates the professional action of the policemen, who use violence to stop this aggressive behavior until Rodney King starts cooperating.

In this fine grained reconstruction, Goodwin shows how Sgt. Duke offers a perceptual field – a coding scheme of the behavior of the policemen, which is in accordance to their professional practice. Hence, they were sentenced not-guilty in the first trial. By highlighting and embedding in the perceptual framework of the profession, some convincing interpretation is constructed and presented as facts, supported by graphical representations of the video-stills.

Goodwin presents – together with his other examples – a detailed study about professional vision. He shows how members of a pro-

fession share a framework of coding schemes, how their practices of seeing are embedded in a community (a phenomenon that has also been addressed with the terms *Denkkollektiv* [Fleck 1981] or *Sehgemeinschaft* [Raab 2008]). However, in Goodwin’s paper, the concept of community of practice that shapes the practices of seeing remains rather abstract and the practices he identifies are apt to fit most processes of seeing.

Few studies exist that focus on vernacular video in a more concrete sense. Most of them are coming from a methodology background or they are only indirectly addressing the process of analysis.¹ First, there is some reflection in education studies on video analysis and how it is used for self-reflection (Hietzge 2008). Furthermore, Laurier, Strelbel and Brown (2008) look at the editing practices of professional movie editors. Finally, detailed studies have been published that deal with the production of video in the process of recording (Macbeth 1999; Mondada 2003; 2005). They also treat video not as a resource, but an object of scrutiny.

In their empirical studies, Tutt and Hindmarsh look into a data sessions conducted by social scientists, especially focusing on how the participants interact with each other and actively generate a shared understanding of what is going on in the video material. In the first paper, they draw attention on the side-work necessary to coordinate distributed research teams

¹ See also Antaki et al. (2008) for an reflection on conversation analysis data sessions.

(Tutt et al. 2007) to focus on an element visible on screen and the interactive coordination of the highlighting of this element. Based on the concept of re-enactment by Sidnell (2006), the authors show how gestures are used by the participants to render the phenomena on the screen visible to the participants in the data session, who are usually sitting in front of a display or screen. Data sessions are a typical form of interaction (in social science, but also in other fields as well), where a small number of participants is discussing some data fragments and interactively producing interpretations.² In the recent paper, Tutt and Hindmarsh (2011) show in detail how speech and gestures as part of the work of interpretation are inter-related, how the person that is highlighting a certain element on screen is creating a gesture space, which binds his co-participants attention and then is able to connect the action visible on screen with the printed transcript. This gesture space is important, because it can be returned to in subsequent moments of the interaction.

Re-enactments are not the only participants’ possibility to solve the interactional problem depicted here. They can apply other resources to communicate their visual knowledge, but we find them in most cases of vernacular video analysis conducted in a discursive manner and with participants co-present. [There are forms of video analysis conducted by a sole analyst in front of a computer – when the pro-

² See Heath et al. (2010:156) for some methodical characterization of data sessions.

cess is routinized. Here, situations where the sole analyst is confronted with ambiguity and problems will exhibit the features of the interpretation to us. But, for us, the discursive, interactional forms of this practice are a better starting point, for understanding this process.] Re-enactments are combined with other communicative methods that we can find in situations of video-analysis. If we want to understand how the processes of data analysis are accomplished in practice, the study of those methods of producing meaning is the relevant step in the research process. Further empirical work should start by collecting those vernacular methods. In my empirical example in the subsequent part of the paper, I am going to focus on the role of enactments, as well as some other methods that need to be addressed in similar fashion.

I now have shown that the production of meaning from video data is not just a cognitive process, but an interactional phenomenon that is produced by body movement, gaze and orientation. The participants of data sessions monitor not only the screen, but also each other, create shared spaces of understanding and fill those with meaning. The following example is taken from a specific public situation of analysis that adds some features of a presentation to it. As the analysis is a very complex process of instruction, understanding and interpretation, the presenters are encouraged to explicitly account for some of the basic interactional processes that are even more clearly articulated in this public demonstration. In situations of "backstage" work on the videos, those actions

can be routinized and do not have to be accounted for participants not acquainted with the specific data. It will be useful to see how actors deal with uncertainty, when encountering new visual recordings.

Example

The following case shows the presentation of an analysis of an incident of violence. It is taking part at a conference³ that is hosted by a hacker and privacy activist group. Two speakers (in my transcript called Peter and Jörg) present and analyze video recordings of a recorded situation. They have prepared the recordings. The actual material analysis at the conference is preceded by a presentation that provided the audience with some background information and legal evaluations. During the analysis that is performed in a demonstrative way, results are presented, rather than new findings generated *in situ*. I will use this datum to highlight some of the basic work units of collaborative video analysis. Before looking at the actual analysis, there will be a look at the "object" of the speaker's investigation.

The data that is taken to vernacular video analysis by the actors has been recorded at a demonstration against surveillance of the public in Berlin in 2009. The accusation was made, was that a young man, who had asked a member of police for his identification number, had been beaten and injured in a case of non-justified

³ The public presentation was filmed, streamed and is available for download by the organizers of the conference.

violence.⁴ The data, covering this event, consist of synchronized recordings from different sources, namely police cameras, as well as camera phones and other camcorders used by surrounding participants of the demonstration.

The speakers (Peter and Jörg, they will be shown later in Figure 2) are responsible for the preparation of the material and state that they asked a production company to produce a synchronized split-screen⁵ version of the recordings that have been made at the demonstration. As one can also see in the video, the recordings cut together have been taped with a variety of handheld video-cameras (some cell-phones with video cameras or similar devices), by participants of the demonstrations.



Figure 1. The analyzed video segment. Source: recorded demonstration against surveillance of the public in Berlin, 2009.

⁴ When writing about this analysis, I am not taking a position whether the accusations made by the analysts are right or wrong. Here, I am not judging the case, but rather interested in how the process of interpretation is demonstrated.

⁵ Mondada (2009) has shown the relevance of the split screen for a variety of professions that use video (and analysed its use in TV debates, where she is interested in real-time, on-air interaction between a number of speakers, that are cut together).

Figure 1 shows the screen in more detail. We can see four different perspectives on the action going on, on the top left side one recording made by a police team combined with the other recordings. The perspectives are "unified" by the time code on the top right corner, which highlights the fact, also mentioned by the presenters, that the videos are synchronized in one objective timeframe, that the viewer can align with.

I want to highlight here, that there already exists the assumption, that the different "subjective" recordings of a situation combined to a multi-perspectival stream of visual information can give us access to one "real" event. This way of understanding the world via the video as an "objective, but incomplete" representation reveals the everyday understanding of interpretation of the visual: real events happen, but some things might be hidden and not be registered by the camera's eye. The assumption is: if one combines all perspectives available in the situation, one is able to understand the real events by seeing it. Not surprisingly, the totality of the video material is presented as fulfilling the role of a witness, but not one that gives an individual narrative account; instead it surrounds the event and looks at it from different visual perspectives that seem to form an almighty observer (interestingly created not by police, but by the video-activists). The way to demonstrate this case of violence is not by gathering all witnesses present at the demonstration or their narrative accounts, but rather just by collecting their recordings. Of course, the speakers at the conference selected earlier on, while preparing the data, which material is relevant, and which is excluded from the screen. Those

processes might also have been issues of discussions or just part of routinized visual expertise.

The multiperspectivity produced here is synchronized in time, but not in space. It is very hard for us to tell the spatial organization of the event or the relation of the objects by each other. The spatial self-allocation of the viewer is produced *in situ* by the speakers (being visible in Transcript 1, for example, in line 9, “up front”).

Let us now look at the analysis going on at the conference: the transcript (Transcript 1) I am presenting here focuses on the speakers analyzing this data and consists of a short excerpt, some minutes into the analysis, where the proceedings of the beating are explained. Both speakers are speaking in turns (and sometimes in a dialogical manner). They are standing on a stage, visible to a big audience (about 200) in front of a projector-screen. Peter (on the left) is pointing towards a relevant detail in the video (I will elaborate on his actions later on). The situation is quite similar to the constellation one can find in PowerPoint presentations (cf. Schnettler and Knoblauch 2007; Knoblauch 2008). Actually, it is preceded by one.

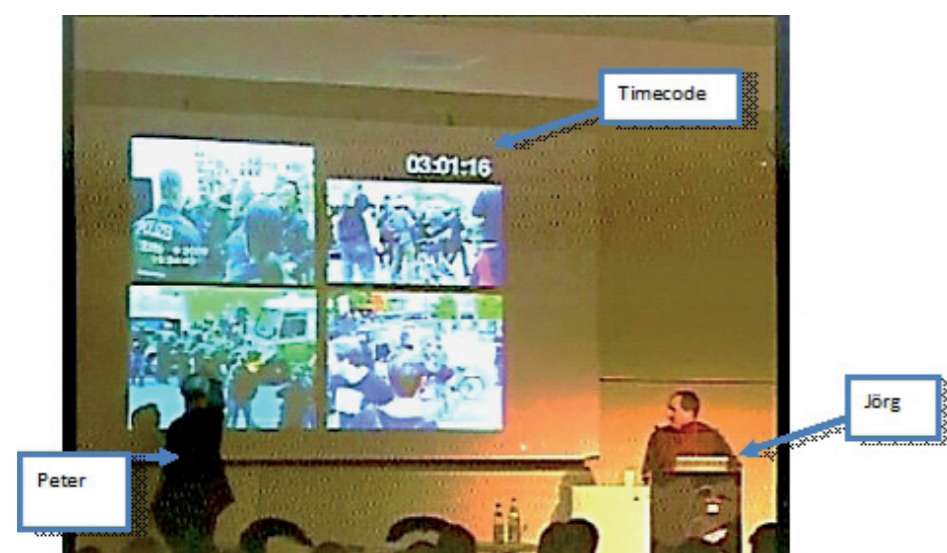


Figure 2. Peter and Jörg, speakers. Source: recorded conference.

- 1 Peter This is the one (-) who we always call the BEATER (.)
Da ist jetzt sozusagen der, (-) den wir immer SCHLÄGER nennen (.)
- 2 Laughter and clapping in the audience,
- 3 Peter This is the one
[No, please, stop it, what's happening here is not nice
Das ist auch der
[Ne hört mal auf, das ist ja nicht schön was da passiert
- 4 Jörg is starting the video, which plays for 2s
- 5 Peter [stop it, please
[halt doch mal an
- 6 Jörg Mhm
Mhm
- 7 Peter This is the one who later on (.) BEATS and then there is another one, who we call the RIPPER. The beater is also the one that,
Das ist auch der der später (.) SCHLÄGT und dann gibt es noch einen den wir REISSER nennen. Der Schläger ist aber auch der,
- 8 Peter that you have seen, as he pushed the cyclist forward
den ihr gesehen habt, wie er den Fahrradfahrer nach vorne gestoßen hat
- 9 Peter the beater is the one that guided the cyclist up front in the scene earlier on.
der Schläger ist auch der, der den Fahrradfahrer in der Szene davor nach vorne geleitet hat.

Transcript 1. Simplified transcript of the analysis. Source: self-elaboration.

The transcript shows a short fragment from the beginning of the analysis. The scene has been introduced, the split screen has been explained as a synchronized documentation and the sources of the videos have been discussed by the speakers – Peter and Jörg. Jörg controls the playback of the recording with his notebook and at the depicted situation is just playing a short fragment, which is paused shortly before the transcribed sequence begins. Peter introduces the following “scene” on screen as “not without relevance” (*nicht ganz unwichtig*).

This sequence of the analysis⁶ is very important for our understanding as well because a coding scheme for the further understanding of the analysis is produced: speaker Peter (on the left hand side of Figure 2) is pointing at the bottom left image of the split screen and identifies a relevant actor. This actor is not just spotted, but (with some irony, emphatically pronounced, which causes the laughter) called “the BEATER” (line 1). The cause for this name is explained (because he is going to beat later on); the actor is introduced by his future action, which is anticipated by the analysis. As this is a case of demonstration, the speakers have looked at their video data before and are bringing their knowledge (typifications, relevancies) into the situation of analyzing this still. The coding scheme in this video is based on a narration containing some protagonists acting in a typical manner – there is a victim (the cyclist), and some of the policemen planning their aggressive actions.

⁶ One has to differentiate between a sequence of analysis, which is produced by the actors (e.g., one argument in the analysis) we observe and the scene or fragment they are observing (one video-clip the actors are playing).

The analysis is conducted on the still – the video is playing, and as soon as a relevant element comes into play it is stopped (as in line 5), where Peter asks Jörg to stop it. This interplay between stopping and commenting and showing a moved image is used to construct continuity in the story. Certain elements, such as the introduction of the relevant actors, background information, spatial arrangement, following or preceding actions are integrated in the pauses. The video – with its time code and continuous play (there is no major jumping backward in the playtime of the video) – is used for the production of a shared continuity.

The continuity of the story consists of the production of a shared time (via visible time code and storytelling), whereas the pauses are used to elaborate on the analytic details: they are supported by the highlighting and inclusion of specific movements on screen that are attributed with a specific meaning. Peter and Jörg have in another sequence, just before our transcript, interpreted that the police officers are “preparing for a criminal act against the cyclist.” To support this severe argument, they produce many arguments that are constantly reintegrated into their story (e.g., in a later episode they are going to tell that the cyclist had asked one of the police officers for his identification number, which is interpreted as a motive for the aggression).

For this complex task, a communicative problem remains: How are the “story,” the participants and the visible elements on screen brought together to a univocal interpretation of this? The

story on screen is not only told, but brought into the physical sphere for the co-participants by bodily action. Just as analyzed by Tutt and Hindmarsh, the ongoing action on screen – is brought back into the situation by re-enactment. Figure 3 consists of combined screenshots of Peter – who does not only tell about the action that the audience has seen before, but he is physically reproducing the movement of pushing and guiding. The form of representation here seems to highlight the role of the gesture, but I want to highlight that not only the gesture is necessary, but rather the orchestration with the other modalities and also the video replay technology. Peter is not just replaying the action, but more importantly, he is "orchestrating" (Schnetler 2006) spatial orientation (forward, upfront) with the typification of the actors and integrating them into the storyline.

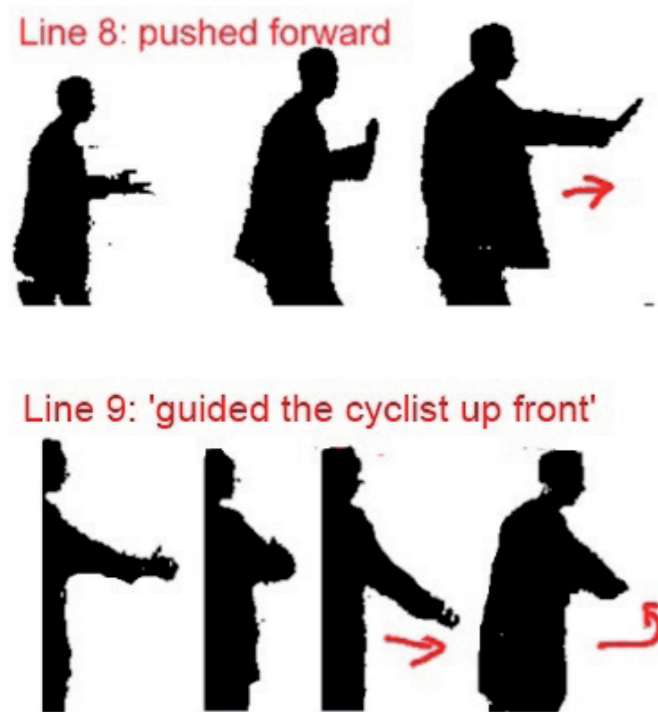


Figure 3. Spatial orientation and re-enactment. *Source: self-elaboration.*

Conclusion

The data analysis (or here: presentation of such) is a situated process in which an event is discursively reconstructed. With the help of a specific technology that allows for combining perspectives, pausing and replaying a temporal and spatial allocation for the participants (and here: audience) is produced. Then a specific storyline with matching roles for relevant observed agents is constructed in which the visible forms of conduct are integrated. The specific elements of knowledge that are actively generated in this communicative process are repeatedly connected to each other. Certain typical movements and gestures are highlighted, interpreted and integrated in to the coding scheme that is produced via the production of a storyline.

Looking at the practices of interpretation and the presentation, I have shown that in this specific case especially the visual conduct was in the focus of attention and the spoken word audible on tape is neglected, focus is given to movement, typical gestures and mimic. Surely, this is due to the specific kind of data the participants are interested in, where the verbal interaction is hard to understand because of the multiperspectivity and the noise at the demonstration; however, it is quite interesting that a live commentary which is produced by the policemen seems not systematically integrated in the analysis. The examples I have presented here are – as mentioned above – taken from a specific kind of visual presentation in which already established interpretations are pre-

sented, made understandable and "proved." One can distinguish such situations with the aim to convince the audience with a certain construction of the event (similar to the Rodney King case) from those where an interpretation is not yet available. In data sessions with new data, this process is more complicated and might contain some other forms that actors actively produce (like guessing, showing uncertainty, communicating imaginary objects, etc.), but the ones I was able to show here were very explicit and can guide attention towards the interactive dimension of interpreting video.

Further work is necessary in systematically distinguishing details – which will also allow us to learn about the communicative character of visual interpretation work (maybe beyond the specific case of "vernacular video analysis"). I have chosen this example because it entails some of the characteristics, presented in interpretation/data sessions of human conduct in a number of fields: alignment in time and space, building a story or theory as coding scheme, selecting and highlighting certain movements, identifying them as actors, interpreting action as meaningful and connecting those elements to one coherent interpretation.

There are further elements one could analyze in this data: What kind of "theory" do the analysts develop on the on-going actions? What are their assumptions they are presenting and how is evidence further on constructed? Which other elements do they take and how do they relate them to their storyline? How do they deal with ambiguities and contradictions? This

details cannot be discussed in this short paper – but are part of an on-going ethnographic study.

Video analysis is a form of re-constructing meaning using audio visual data. As many social scientists do use visual data for research aims, the methodical basis has been laid. However, when looking at the practices of video in more vernacular fields, there are only a few studies that really show how video is used in practice. How does the availability of this technology and the spread of the practice of analyzing video data for a variety of ends generate social reality?

I argue that there are some specificities of video-analytic practices that go beyond the general practices identified by Goodwin, but only more empirical work will show how video is used in a variety of vernacular fields, and in how far the specific expert knowledge available in a field relates to the practices of professional vision. Looking at video-interpretation, I want to highlight that the production of those interpretations is not only a cognitive, but an interactive process based on bodily re-enactment and pointing, as well as orchestration with spoken language. Visual knowledge, then, is not only visual, but integrated into a multimodal process of instruction and understanding.

References

- Antaki, Charles, Michaela Biazzì, Anette Nissen et al. 2008. "Accounting for moral judgments in academic talk: The case of a conversation analysis data session." *Text & Talk – An Interdisciplinary Journal of Language, Discourse & Communication Studies* 28(1):1-30.
- Bateson, Gregory. 1958. "Language and Psychotherapy: Frieda Fromm-Reichmann's last project." *Psychiatry* 21(21):96-100.
- Bateson, Gregory and Margaret Mead. 1942. *Balinese Character. A Photographic Analysis*. New York: New York Academy of Sciences.
- Birdwhistell, Ray L. 1952. *Introduction to Kinesics. An Annotation System for the Analysis of Body Motion and Gesture*. Louisville, KY: University of Louisville.
- Brookman, Philip. 2010. *Eadweard Muybridge*. London: Tate Publishing.
- Burri, Regula and Joseph Dumit. 2008. "Social Studies of Scientific Imaging and Visualization." Pp. 297-317 in *The Handbook of Science and Technology Studies, 3rd Edition*, edited by E. J. Hackett, O. Amsterdamska, M. Lynch, et al. Cambridge, London: MIT Press.
- Ekman, Paul and Wallace V. Friesen. 2003. *Unmasking the Face: A Guide to Recognizing Emotions from Facial Expressions*. Cambridge: Malor Books.
- Erickson, Frederick. 2011. "Uses of video in social research: a brief history." *International Journal of Social Research Methodology* 14(3):179-189.
- Fleck, Ludwik. 1981. *Genesis and Development of a Scientific Fact*. Chicago: University of Chicago Press.
- Goodwin, Charles. 1994. "Professional Vision." *American Anthropologist* 96(3):606-633.
- Goodwin, Charles. 2000. "Practices of Seeing: Visual Analysis: An Ethnomethodological Approach." Pp. 157-182 in *Handbook of Visual Analysis*, edited by T. Van Leeuwen, C. Jewitt. London: Sage.
- Greenberg, Joshua H. 2008. *From Betamax to Blockbuster: Video Stores and the Invention of Movies on Video*. Cambridge: The MIT Press.
- Heath, Christian, Jon Hindmarsh and Paul Luff. 2010. *Video in Qualitative Research*. London: Sage.
- Hietzge, Maud. 2008. "Videogestützte Selbstreflexion in der Sportlehrerausbildung – Reaktivität, Akzeptanz und «how to do»." Pp. 295-298 in *Jahresband der Deutschen Vereinigung für Sportwissenschaft (DVS)-Sektion Sportpädagogik*, edited by V. Oesterheld. Hamburg: Czwalina.
- Kissmann, Ulrike T. 2009. *Video Interaction Analysis: Methods and Methodology*. Frankfurt: Lang.
- Knoblauch, Hubert. 2008. "The Performance of Knowledge: Pointing and Knowledge in Powerpoint Presentations." *Cultural Sociology* 2(1):75-97.
- Knoblauch, Hubert, Bernt Schnettler, Jürgen Raab et al., (eds.). 2006. *Video Analysis – Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt am Main: Lang.
- Knoblauch, Hubert, Alejandro Baer, Eric Laurier et al. 2008. "Visual Analysis. New Developments in the Interpretative Analysis of Video and Photography." *Forum Qualitative Social Research* 9(3), Art. 14. Retrieved July 26, 2012 (<http://www.qualitative-research.net/index.php/fqs/article/view/1170/2587>).
- Knoblauch, Hubert, Bernt Schnettler and René Tuma. 2010. "Interpretative Videoanalysen in der Sozialforschung." *Enzyklopädie Erziehungswissenschaft Online*, edited by S. Maschke, L. Stecher [www.erzwissonline.de]. Weinheim: Juventa.
- Laurier, Eric, Ignaz Strebel and Barry Brown. 2008. "Video Analysis: Lessons from Professional Video Editing Practice." *Forum Qualitative Social Research* 9(3). Retrieved July 14, 2012 (<http://www.qualitative-research.net/index.php/fqs/article/view/1168>).
- Macbeth, Douglas. 1999. "Glances, trances, and their relevance for a visual sociology." Pp. 135-170 in *Media studies: Ethnomethodological approaches*, edited by P. L. Jalbert. Lanham: University Presses of America.
- Mittenecker, Erich. 1987. *Video in der Psychologie. Methoden und Anwendungsbeispiele in Forschung und Praxis*. Bern: Huber.
- Mondada, Lorenza. 2003. "Working with video: how surgeons produce video records of their actions." *Visual Studies* 18(1):58-72.
- Mondada, Lorenza. 2005. "Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis." Pp. 51-67 in *Video-Analysis*, edited by H. Knoblauch, B. Schnettler, J. Raab, et al. Frankfurt am Main: Lang.
- Mondada, Lorenza. 2009. "Video Recording Practices and the Reflexive Constitution of the Interactional Order: Some Systematic Uses of the Split-Screen Technique." *Human Studies* 32(1):67-99.
- Raab, Jürgen. 2008. *Visuelle Wissenssoziologie. Theoretische Konzeption und materiale Analysen*. Konstanz: UVK.
- Reichert, Ramón. 2007. *Im Kino der Humanwissenschaften: Studien zur Medialisierung wissenschaftlichen Wissens*. Bielefeld: transcript.
- Ruby, Jay. 2000. *Picturing Culture: Explorations of Film and Anthropology*. Chicago: University of Chicago Press.
- Schnettler, Bernt. 2006. "Orchestrating Bullet Lists and Commentaries. A Video Performance Analysis of Computer Supported Presentations." Pp. 155-168 in *Video Analysis – Methodology and Methods. Qualitative Audiovisual Data Analysis in Sociology*, edited by H. Knoblauch, B. Schnettler, J. Raab, et al. Frankfurt am Main: Lang.
- Schnettler, Bernt. 2007. "Auf dem Weg zu einer Soziologie visuellen Wissens." *Sozialer Sinn* 8(2):189-210.
- Schnettler, Bernt and Hubert Knoblauch. 2007. *Powerpoint-Präsentationen: Neue Formen der gesellschaftlichen Kommunikation von Wissen*. Konstanz: UVK.
- Sidnell, Jack. 2006. "Coordinating Gesture, Talk, and Gaze in Reenactments." *Research on Language & Social Interaction* 39(4):377-409.
- Thiel, Thomas. 2003. "Film und Videotechnik in der Psychologie. Eine erkenntnistheoretische Analyse mit Jean Piaget, Anwendungsbeispiele aus der Kleinkindforschung und ein historischer Rückblick auf Kurt Lewin und Arnold Gsell." Pp. 649-708 in *Handbuch der Kleinkindforschung*, edited by H. Keller. Bern, Göttingen, Toronto, Seattle: Hans Huber.
- Tutt, Dylan and Jon Hindmarsh. 2011. "Reenactments at Work: Demonstrating Conduct in Data Sessions." *Research on Language & Social Interaction* 44(3):211-236.
- Tutt, Dylan, Jon Hindmarsh, Muneeb Shaikat et al. 2007. "The distributed work of local action: Interaction amongst virtually collocated research teams." Pp. 199-218 in *ECSCW 2007*, edited by L. J. Bannon, I. Wagner, C. Gutwin, et al. London: Springer.
- Zielinski, Siegfried. 1986. *Zur Geschichte des Videorecorders*. Berlin: Wissenschaftsverlag Spiess.
- Tuma, René. 2012. "The (Re)Construction of Human Conduct: «Vernacular Video Analysis»." *Qualitative Sociology Review* 8(2):152-163. Retrieved Month, Year (http://www.qualitativesociologyreview.org/ENG/archive_eng.php).