

Hubert Knoblauch/Bernt Schnettler
Jürgen Raab/Hans-Georg Soeffner
(eds.)

Video Analysis: Methodology and Methods

Qualitative Audiovisual Data Analysis in Sociology

2nd, revised edition



PETER LANG

Frankfurt am Main · Berlin · Bern · Bruxelles · New York · Oxford · Wien

Bibliographic Information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the internet at <<http://www.d-nb.de>>.

Cover image by
Bernardo Fernández.

Reproduces with kind permission of the artist.

SW
2009
P77



K
ISBN 978-3-631-59239-7

© Peter Lang GmbH
Internationaler Verlag der Wissenschaften
Frankfurt am Main 2006
2nd, revised edition 2009
All rights reserved.

All parts of this publication are protected by copyright. Any utilisation outside the strict limits of the copyright law, without the permission of the publisher, is forbidden and liable to prosecution. This applies in particular to reproductions, translations, microfilming, and storage and processing in electronic retrieval systems.

Printed in Germany 1 2 3 4 5 7

www.peterlang.de

Contents

Introduction

Hubert Knoblauch, Bernt Schmettler & Jürgen Raab
Video-Analysis. Methodological Aspects of Interpretive Audiovisual Analysis in Social Research 9

Methodologies of Video Analysis

Thomas Luckmann
Some Remarks on Scores in Multimodal Sequential Analysis 29

Christian Heath & Paul Luff
Video Analysis and Organisational Practice 35

Lorenza Mondada
Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis 51

Hubert Knoblauch
Videography. Focused Ethnography and Video Analysis 69

Jürgen Raab & Dirk Tänzler
Video Hermeneutics 85

Research Fields of Video Analysis

Dirk vom Lehn & Christian Heath
Discovering Exhibits: Video-Based Studies of Interaction in Museums and Science Centres 101

Cornelius Schubert
Video Analysis of Practice and the Practice of Video Analysis. Selecting field and focus in videography 115

Anssi Peräkylä & Johanna Ruusuvuori
Facial Expression in an Assessment 127

Monika Wagner-Willi
On the Multidimensional Analysis of Video-Data. Documentary Interpretation of Interaction in Schools 143

Bernt Schnettler
Orchestrating Bullet Lists and Commentaries. A Video Performance
Analysis of Computer Supported Presentations..... 155

Practices of Video Analysis

Elisabeth Mohn
Permanent Work on Gazes. Video Ethnography as an Alternative
Methodology..... 173

Eric Laurier & Chris Philo
Natural Problems of Naturalistic Video Data 183

Sigrid Schmid
Video Analysis in Qualitative Market Research – from Viscous Reality to
Catchy Footage 193

Epilogue

Hans-Georg Soeffner
Visual Sociology on the Basis of ‘Visual Concentration’ 209

Contributors to this volume 219

Introduction

Hubert Knoblauch, Bernd Schmettler & Jürgen Raab

Video-Analysis

Methodological Aspects of Interpretive Audiovisual Analysis in Social Research

I

In recent years, we have witnessed the proliferation of an increasingly sophisticated new instrument of data collection: Video camcorders. Camcorders do not only allow for a rich recording of social processes. They also provide and produce a new kind of data for sociology. In fact, some authors believe to be able to discern a “video revolution”: the effects of this “microscope of interaction” are expected to be as profound as was the invention of the tape recorder, which gave rise to new research disciplines such as conversation analysis.

In fact, video is much more widely used nowadays in the most diverse branches of society than the tape recorder ever was. Video-art, wedding videos, holiday videos and the huge variety of usages of video on the internet demonstrate to anyone and everyone that video has become a medium that pervades our everyday life. An ever-increasing role is played by video-mediated forms of communication, such as video-conferences (Finn, Sellen & Wilbur 1997). It is quite likely that the dissemination of UMTS will also lead to a more wide-spread use of mobile video mediated communication and video-messaging. Finally, video surveillance technologies have become an accepted part of our daily lives (Fyfe 1999, Fiske 1998).

As accepted and broadly used as camcorders and video records may be in all institutional spheres as well as in private life, the methodological discussion of their use in scholarly studies is greatly underdeveloped. As a medium used by the people themselves, video deserve much closer attention than we are able to pay them in this book. If we, however, look at the science of society – sociology (and, for that matter, other social scientific disciplines) – we discern a wide disregard for video. Whereas text-centred approaches have been subject to innumerable methodological reflections and methodical designs, video has neither as a method of data collection nor as a medium used by the members of our society been able to attract much attention from sociologists and other students of society and culture.

It is for this reason that we would like to take this opportunity to tackle the task of presenting methodologies for the analysis of video. By this we do not mean methodologies for the use of visual data, for this has already been the subject of many books (cf. Banks & Murphy 1997, Davies 1999, Emmison & Smith 2000, Pink 2001).

Instead, we are interested in methodologies that address questions related specifically to analytical work with video recordings.

Very early on, the advantages of video as an observational technique proved to be quite obvious (cf. Gottdiener 1979, Grimshaw 1982, Heath 1986). Compared to observations made by the naked human eye, video recordings appear more detailed, more complete and more accurate. In a technical sense, they are more reliable since they allow data analysis independent of the person who collected the data. However, despite the fact that video now is widely used in sociology and the social sciences, there have been but very few attempts to discuss the methodology of working with this medium as an instrument of data collection and analysis. No doubt, debates on visibility, visual culture and visualisation abound – also in the social sciences. Nowadays, there is a huge amount of criticism at the level of epistemology. Anyone interested in the field will discover flourishing debates on the cultural meaning of video-clips of Madonna's pop songs or the epistemological question as to the hows and whys of the picture's betrayal of the viewer. However, few are the scholars who actually address the question of what to do in case one dares not just to talk about epistemology, but instead to use the medium and work empirically with the data produced within its various forms (Jordan & Henderson 1995, Heath 1986, Lomax & Casey 1998, Heath 1997, for classroom interaction cf. Aufschnaiter & Welzel 2001).

By publishing this volume, we wish to change this situation, at least to some degree. The goal of this book is to provide ways in which videos can be analysed sociologically. The book, then, is an attempt to gather a number of researchers familiar with video analysis in order to focus on, scrutinise and clarify the crucial methodological issues in doing video analysis. The questions we would like to tackle are: what are the central features of video data; what kinds of video data can be distinguished; and particularly how should we analyse and interpret video data? In trying to answer these questions, the book will provide support for all those who are planning to use video as an instrument of data collection and analysis.

II

When we speak of video analysis, it should be stressed that we are not referring to any and all kinds of work with video. To the contrary, there are a number of qualifications to the kind of studies represented in this volume which must be named in addition to all those features mentioned in the papers. First, it will become quite obvious that we have limited the range of studies presented to *social scientific analyses* of video data. People, their actions and the structures constructed by these actions lie at the heart of what is of interest to these studies. Within the social science framework, a variety of disciplines will be represented: sociology, anthropology, linguistics and education – as well as a number of researchers who would locate themselves across these disciplines or

in fields in which their studies are being applied (such as architecture, city planning or design). For the sake of brevity and for other contingent reasons, we have to concede that a number of disciplines are not represented in which video analysis has gained some importance, such as the psychology of perception or the visual arts.

The range of disciplines and the kind of video analysis portrayed in this volume share a second feature. Whoever scans the contributions in this book will soon discover that they seem to share a similar topic. Across the variety of fields, most of the studies focus on what one would call activities and interaction. Be they studies of the use of high technology and workplace settings, be they studies of people visiting museum, science studies or classroom investigations etc. – all of them to focus on *visual conduct* in general and on *interaction* in particular. It is the focus on the *audiovisual aspects of people in action* which constitutes the central subject of these video analyses. In more theoretical terms, one could say that the field of video studies is circumscribed by what Erving Goffman called the 'interaction order', i.e. the area of action in which people act in visual co-presence – a co-presence which can be captured by the camera. And since what people do covers a huge range of areas, the potential topics of video analysis is almost endless.

As varied as the topics may be, the manners in which the authors approach their topics are just as distinct. Although video analysis initially privileged experimental settings and studios, the kinds of analyses included here turn to what has come to be called "*natural data*". Of course, natural data does not resemble the data found by natural scientists; since all video analysts agree in the interpretive character of their data, there should be no misunderstanding of natural data in this sense. Instead, by natural data we mean that the recordings are made in situations affected as little as possible by the researchers (Silverman 2005). Natural data refers to data collected when the people studied act, behave and go about their business as they would if there were no social scientists observing or taping them. There is no doubt that the very presence of video technology may exert some influence on the situation that is being recorded, an influence commonly labelled 'reactivity'. In fact, this issue is addressed in this volume. Nevertheless, many studies show that the effect of video becomes negligible in most situations after a certain phase of habituation. The stress on the naturalness of data should, however, not be understood as a total neglect of other kind of situations. Interviews or even experiments may also be subjected to video analyses, the general assumption being that they are not as a result taken to represent something else (i.e. what is talked about in the interview), but only as what they are: interviews or experiments. In general, however, video analyses turn to more profane situations: people at work, people in the museum, people sitting in a café etc. It is, by the way, this orientation towards "natural situations" that leads video analysts to sympathise strongly with ethnography, particularly the kind of ethnography which turns towards encounters, social situations and performances as championed by Erving Goffman (1961, 1967, 1971). In order to distinguish this ethnographically oriented video analysis from other standardised forms of video analysis, it seems therefore quite reasonable to apply to it the term 'videography' (cf. Knoblauch, this volume).

However, although the “naturalness” of the data is a goal towards which video analysts in general strive, it would be misleading to assume that there is only one sort of data for video analysis. Rather, there is a whole array of what may be called “data sorts” produced by video data collection. There are two reasons for this variety: first, because people in “natural situations” may themselves use video recording technology, they provide video researchers with various sorts of videos, such as weddings videos, videos from other festive occasions or bits and pieces of their everyday life. Second, researchers may produce videos in differing ways. They may, for example, ask the actors themselves to portray their everyday life by means of the video, e.g. by producing video diaries¹; they may actively use the camera as an instrument of visual construction of data or they may edit the video data in various ways which are now much more readily accessible. On these grounds, we would suggest distinguishing between various sorts of video data. By sorts of video data, we refer to the ways in which the data are constructed (cf. Knoblauch 2003: chap. III). Some sorts of video data are sketched on the diagram below. The ways in which data is constructed may be distinguished in two dimensions: on the one hand, the data are manipulated through various technical procedures. No doubt, the technical recording itself may be considered a decisive form of manipulation. However, whereas different technologies (Super 8, V 8, digital video etc.) produce almost the same results, the differing technologies allow for an additional set of manipulations: beginning with repeating, slow motion and single frame, these include ways of selection, highlighting, enlargements etc. We subsume all these forms under the label “record”. Secondly, videos may be distinguished by the way they address the situation. Whereas some just try to “copy” what has been visualised, others attempt to make something seen which is not happening without their influence. Wedding guests wish to see the newlyweds kissing each other in front of the camera; the experimenter wishes that the subjects shake hands, the film maker wishes the actors to hit each other. This level of manipulating the situation for the sake of what may be seen on the video by the recipients we call ‘recipient design’. Within these two dimensions we can locate a number of data sorts: video-diaries, weddings videos, “natural videos” etc.

The studies represented in this volume share an additional common feature. Whereas in a number of fields, e.g. in psychology or in engineering, we find a strong tendency to standardise, even automatise data analysis (Mittenecker 1987, Koch & Zumbach 2002), the contributors of this volume propose a rather different methodology. It is not that they oppose standardisation or automatisisation in general. However, they all share the conviction that it is definitely premature to approach audiovisually

1 Thus Holliday (2000) asked subjects to produce ‘auto-ethnographic’ videos in order to show how they organize their daily lives. In a similar way, in Anthropology, for example, indigenous people have been asked to use the video in order to preserve their “native” perspective (cf. Ruby 2000).

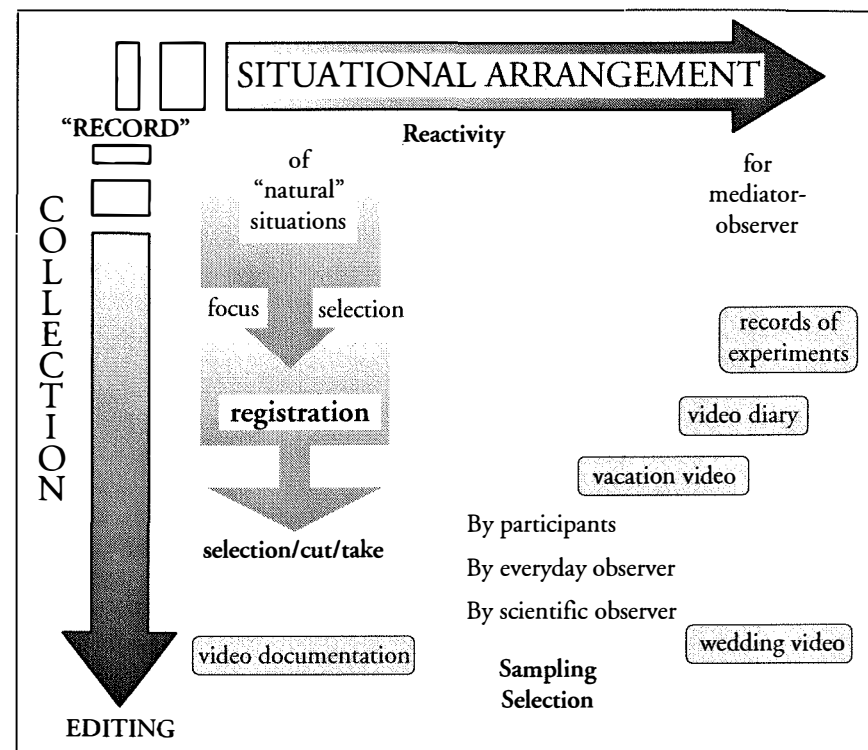


Fig. 1: Methodological Dimensions of Video-Data Collection

recorded data by means of standardised coding procedures. Instead, all of the methods suggested here can be said to relate to what is commonly called *non-standardised, qualitative* or, to be more exact, *interpretive social research*.² They share the assumption that the world in which people act is a world of meanings and that, therefore, research on people in action must account for the meaning of these actions. Yet it would be utterly misleading to assume that the volume’s methodological orientation is monolithic. Within the field of interpretive or qualitative video analysis, there is still quite a variety of approaches. The volume tries to collect at least the most prominent of them. Ethnomethodology and conversation analysis represent, of course, major fields, as do genre analysis, grounded theory and sociological hermeneutics.

2 To Pink (2001), reflexivity is the major feature of visual anthropology in general and video studies in particular. In our view, reflexivity is subordinated to the demand for interpretation – a demand which goes back to founding fathers of interpretive social sciences such as Weber and Schutz.

Thus, the papers as a group share a series of topics which are crucial for the current state of video analysis. In addition to the common orientation as social scientific, naturalistic and interpretive studies of social interaction, all papers stress that *sequentiality* is fundamental to video analysis. Although sequentiality can mean various things, (particularly between the hermeneutic notion and the rather conversation-analytical one), the parallel between the sequentiality of the medium and the sequentiality of social activities is fundamental to video analysis. Since all approaches are interpretive, the analyses build in one way or the other on what may be called “ethno-hermeneutics”. They also share the methodological conviction that interpretive analysis of video-data requires more than “visual empathy” combined with a mainly descriptive “structured microanalysis” as Denzin (2000) suggests.

III

There is no doubt that the book cannot at once solve all the problems of video analysis. To the contrary, the papers presented here permit us to identify a series of issues that urgently need to be tackled. First, the problem of *complexity*: the relative neglect of video in the social sciences is sometimes attributed to its complexity and abundance. A few minutes of recording produce a large quantity of visual, kinaesthetic, and acoustic data that must be transcribed and prepared for analysis. Video data is certainly among the most complex data in social scientific empirical research. It is multi-sensual and sequentially ordered, enclosing both diachronic and synchronic elements, e.g. speech and visual conduct, gesture, mimic expressions, representation of artefacts and the structure of the environment, as well as signs and symbols. Moreover, it represents aspects related to recording activity itself, such as the angle and the focus of the camera, the cuts, and other elements pertaining to the activity of filming and editing. Hence, video recording generates an extraordinary abundance of data, confronting the researcher with the problems of data management, retrieval and selection. This may not only cause the problem of data overload, but also raises the question of how to select sequences appropriate for further scrutiny. It might also be the case that the quality of the recordings may be detrimental to analytic purposes. There may be interesting parts of video that can not be selected for further scrutiny due to, for example, recording problems (wrong perspective, defect in recording, people running through the image, etc.). Beyond such obvious practical restrictions, the methodological problem of what constitutes the unit of analysis and how to assure a balance between time-consuming microanalysis and an overview over the whole data corpus remain open questions for future methodological debates.

The second problem to be tackled urgently is the *technological challenge*. The role of technology should be taken into consideration to a much stronger degree than we can do here. The very fact that the methodology is heavily based on a technology subjects

it to future technological developments. This does not only raise the question mentioned above of what impact the technology may have on social scientific video analysis (and *vice versa*). Video confronts the researcher with a number of technical and material challenges. Some of them concern the implementation of camera, microphones, software etc. This technical part is still underestimated in the methodological discussion. Even if technology may not be considered an “autonomous actor” (Rammert & Schulz-Schaeffer 2002), the employed artefacts definitely exert at least some influence upon the course of action in the research process. Without doubt, the instruments change the way in which we collect, construct, analyse and interpret our data. Methodological considerations rarely reflect this material issue because we are used to discussing methodology in much more abstract terms. Hence, we may ask in which ways the instruments interfere with our analytical work. This question is especially pertinent for video analysis, which, compared to other qualitative methods, requires quite a lot of technology. Indeed, it may represent one of the most expensive and intricate ways to conduct qualitative research. Fortunately, equipment has become much cheaper and easier to handle in the last few years. Today, filming does not cost us 30,000 German Marks as it did when social psychologist Kurt Lewin started using films in the 1930's (Thiel 2003). Nonetheless, researchers still must purchase camcorders, tapes, tripods, microphones, etc. for the purpose of recording videos. In addition, analysing video data requires intelligent storage and cataloguing systems for raw data, powerful computer hardware and a series of software tools to digitalize, transcribe and analyze data and to present research results. Due to miniaturization and popularization, a very basic version of video equipment has even become accessible for students. Nevertheless, expenses entailed for basic research equipment (somewhere between equipment available for popular use and that used by television professionals) easily may amount to tens of thousands of Euros – in addition to the space, time and patience required to select the appropriate apparatus and software. Its handling requires also novel technical skills, quite unprecedented in qualitative inquiry. And, unlike other, more conventional forms of qualitative research, e.g. participant observation or interviews, preliminaries and preparation take considerably more time in qualitative video analysis. This may cause a certain delay in the analytical work, as quite extended portions of time are consumed by mere “craftsmanship”. (As a result, qualitative inquiry may become more similar to quantitative research. As in surveys, much work is invested in preparation, providing skills to the coders, handling the data-collections etc.)

Third, the relation between text and image must be clarified. No doubt, the relation between the spoken and the visual is of general epistemological importance. In the case of video analysis, however, this issue exhibits a very practical aspect: the *transcription* of data inscribes in its particular way how the visual is accounted for by the analysis, so that any further development of video analysis will also depend on the way in which data are being transcribed or otherwise made accessible for analysis. Analysis will increasingly be able to draw on visual representation, with the result that written

transcripts may lose their importance to such a degree as to possibly open the way for a “visual mentality” in analysis – a mode of analysing that depends less on the written word than on visualisation and imagination. The ongoing technological changes may also affect the way (and are already now affecting the ways) in which studies are being presented (cf. for example Büscher 2005). However, for the time being, we still rely on the rather conventional forms of transcriptions and frame grabs which are used in this book.³ Consider that transcribing data is not just a preliminary phase of analysis. It forms an essential part of analysis. Transcribing generates observations that are fundamental to analytical inferences. As in research based on natural communicative activities or interviews, the transcription of video data is simply indispensable.

Conversation analysts and linguists have developed a wide array of transcription systems that transform the analytically important aspects of spoken language into textual representations (cf. Dittmar 2002 for a comprehensive overview). Nevertheless, transcription systems for video data still remain in an experimental stage. “There is no general orthography used for the transcription of visual and tactile conduct”. However, “over the years researchers have developed *ad hoc* solutions to locating and characterizing action” (Heath & Hindmarsh 2002: 20?). In this volume, readers will find a variety of approaches for transcribing the visual aspect which, nevertheless, may all be characterized as relatively preliminary. These “ad hoc solutions” are comprised of transcripts consisting basically of detailed description of what occurs in the video. There are also types of transcriptions for the non-verbal aspects and their relation to the verbal behaviour of the participants, ‘conduct score’, and sketches of action sequences or ‘thick interpretative descriptions’ in addition to representations of data that attempt to make use of the visual potential of video data.

Finally, one of the most salient problems is the *legal implications* of video-recording. Like any other form of research, video analysis is subject to legal and ethical restrictions. This concerns questions such as: where are video analysts permitted to film, who is permitted to record social interactions for analytical purposes, which of these images may be stored, analyzed or even used for publication and thereby disclosed to a wider audience. Although there have been intense debates on issues related to video recording in public places, their focus has been primarily on security issues and the questions of infringement on individuals’ right to privacy. To our knowledge, there is no specific regulation for scientific video recordings at the moment.⁴ To assure that some kind of ‘informed consent’ exists seems to be, in the meantime, the most reasonable practical solution, although there may be cases in which this is virtually impossible (e.g. for each single pedestrian in wide-angle shots of public places). In addition,

³ In addition, some of the video recordings analyzed in the different contributions to this book are available at <http://www.tu-berlin.de/fb7/ifs/soziologie/AllgSoz/publikationen.htm>.

⁴ We are grateful to Prof. Dr. Hansjürgen Garstka, the German federal government’s Secretary for Data Security, for his comment on the legal situation in Europe.

unlike for example the case of interview transcripts, anonymisation of moving images is a technically much more demanding task. Consequently, respecting the right to privacy in video analysis is a difficult and as yet unresolved problem, in addition to the legal implications of possible infringements on copy-rights and other rights that may be touched by capturing, recording, analysing, storing or publishing video data of some sort (i.e. the fine distinction the legal systems draws in the field of data protection in general). Legally, the use of video for scholarly purposes of the kind described above oscillates between the individual freedom, which puts particular restrictions on “natural recording” practices, on the one hand, and the freedom of research, which puts no limits on the potential subjects of video recording to the extent that these may be of scientific relevance. Because of the tension between these two extremes, researchers often find themselves caught in a dilemma. We hope that this dilemma will soon find a legal solution.

IV

As mentioned above, the different directions of video research represented in this volume share a number of features: they are social scientific, naturalistic, interpretive studies of visual conduct. As such, they refer, of course, to the long tradition of sociological thinking in general as well as to the study of social action and interaction in particular. In focusing on the realm of the visual, they also draw on the history of visual anthropology and sociology. The era of visual studies was opened at the turn of the last century, when photography and film started to be used within the social sciences (for an example see Breckindrige & Aboth 1910, MacLean 1903, Walker 1915, Woodhead 1904). By means of visual technologies, anthropology developed a visual branch (Collier 1979, cf. Bateson & Mead 1942, Mead 1975, Collier 1967, Collier & Collier 1986). In the form of the much more tenacious development termed visual sociology (Curry 1984, Curry & Clarke 1978, Henney 1986), it focused mainly on photography, and film was used primarily as a means of presenting results than as a datum to be analysed. Famous early examples are A. C. Haddon, Baldwin Spencer or Robert Flaherty who, starting at the turn of the 19th century, used film in order to analyse human conduct. Flaherty, for example, became familiar with the language and culture of the Inuit Eskimo and involved them in the making of his film studies. Another example is “The Ax Fight” by Asch und Chagnon, in which a short, violent fight among the Yanomamo Indians, filmed from a certain distance, is portrayed. The text of the film consists of the comments made by both researcher during the situation filmed (cf. Marks 1995). No doubt, anthropology developed an unprecedented collection of film data which was, as mentioned, mostly used to document reality instead of analysing it (Heider 1976).

The analysis of films as data took another route. As one of the first to use film as a datum for the study of behaviour, Kurt Lewin filmed a behavioural sequence as early as 1923/1924. Lewin analysed this sequence as an example for a behavioural conflict.

Building on Lewin, in 1935 Gesell published a book on "cinema analysis" as a "method for Behavior Study" in which he used frame-to-frame analysis (for more details cf. Thiel 2003). One could consider the famous analyses of Bali dance by Margaret Mead and Gregory Bateson (1942) as a continuation of these studies. In a later study, Bateson and the so called "Palo Alto group" used film in order to analyse interaction between family members. Again, psychologists were included (such as Frieda Fromm-Reichmann and Paul Watzlawick) because the main goal was to investigate if it is interaction that produces the "psychological disturbance" of individual family members. It was also Fromm-Reichmann who initiated the famous project on the "History of the Interview" in which the various modes of interaction were analysed for the very first time (Bateson 1958). Whereas the use of video in psychology increasingly came to focus on what was called "non-verbal behaviour" (cf. the seminal studies by Ekman & Friesen 1969), a parallel development saw the establishment of a marginal stream of studies with employed films to attempt to capture behaviour in a more encompassing and meaningful way. Among these were the studies of Ray Birdwhistell (1952, 1970), who analysed the interplay between nonverbal and verbal behaviour in minute detail, coining the notion of kinesics. (Birdwhistell also has the distinction of being one of Erving Goffman's teachers, who was to become so important for the study of interaction). In a similar vein, Albert Schefflen (1965) analysed the role of posture for the structuring of psychotherapeutic encounters. Until the 1970s, however, these analyses were performed on the basis of film, which was a difficult medium for analysis. Things changed slowly with the introduction, miniaturisation and technical sophistication of video we have witnessed since then. It was particularly among conversation analysts that this medium gained relevance. This might be surprising since, for a long time (and, to some, until now), "hard core" conversation analysis prohibited the use of data of any other sort than audio recordings. On the other hand, the development of conversation analysis was supported by the use of the audio recorder, and the introduction of the camcorder seemed to extend the kind of data collection conversation analysts had been used to. Charles Goodwin was one of the first to use video in the way. He analysed spoken interaction in such a way as to show how visual aspects (particularly gaze) help to bestow order (Goodwin 1986, Goodwin 1981). Erickson and Shultz (1982) used video in their studies of four school counselors in their interview interaction with pupils. Also in the early 1980s, Christian Heath undertook video studies, targeting whole social situations such as medical encounters (Heath 1986). By the late 1970s, Thomas Luckmann and Peter Gross (1977) started a project which used video in order to develop an annotation system for interactions which was compared to a musical score. In a way, this project analyzed what has become to be called multimodality, even if most studies in this volume tackle this issue in a rather holistic way. Whereas this gave rise to a hermeneutic (Bergmann, Luckmann & Soeffner 1993, Raab 2001, 2002) and genre-analytic approach to video (Schnettler 2001, Knoblauch 2004), it was the more ethnomethodological approach of video analy-

sis which became increasingly employed in workplace studies, a field of research preoccupied with interaction at work in high technology settings (cf. Heath, Knoblauch & Luff 2000). It was again Christian Heath and his team who has contributed substantially to this field, as well as Lucy Suchman, Charles Goodwin, and Brigitte Jordan, etc. As far as we can see, it is only within this area that serious reflections on an interpretive methodology of video analysis have been undertaken. Thus, Christian Heath and others have sketched the methodological background of video analysis in several essays (1997) and Suchman & Trigg (1991) have explained the ways in which video contributes to workplace studies. Brigitte Jordan and Austin Henderson (1995) have tried to situate video analysis within the larger framework of interaction analysis. In a similar field of research, the French sociology of work, we even find a whole journal issue devoted to the issues of video analysis and visual sociology (see for example Lacoste 1997).

V

The *papers in this volume* build on this type of video analysis; they are, as we have said, all social scientific, interpretive and naturalistic. As we shall see, their subject is human action and interaction. Despite the similarities, the focuses of the papers varies to some degree, so we have decided to put them in an order that reflects this variation.

The first series of papers focuses on *methodological issues* and address the question how video data may be analysed in a scientific manner. This question is addressed by other papers, too, since it is the common topic of the whole book. The papers in this section directly address this topic and propose analytical methodologies. These papers delineate approaches oriented to conversation analysis, ethnography or hermeneutics and, like THOMAS LUCKMANN in his short paper "Some Remarks on Scores in Multimodal Sequential Analysis", interpretive sociology in general. As he indicates, video provides a very helpful instrument for the analysis of interaction since it, despite all technical transformations, preserves the temporal and sequential structure which is so characteristic of interaction. Nevertheless, video analysis faces some serious problems which may be the reason for what he considers the "backwardness" of this method. It is the integration of the many modes of interaction, particularly the integration of the spoken and the visual, which must be addressed by a successful methodology.

CHRISTIAN HEATH and PAUL LUFF ("Video-Analysis and Organisational Practice") address the methodology of video analysis from a quite unusual and enlightening angle. Instead of sketching the ways in which analysis that meets scholarly standards should be conducted, the authors demonstrate very lucidly how video is analysed by lay persons in our societies. In treating actors whose professions require that they watch and on this basis interpret the behaviour of other actors as represented on video, they show how operators in undergrounds, personnel in surveillance centres and mem-

bers of similar professions act as (sometimes quite sophisticated) “lay sociologists” who must make sense of conduct and interactions. This sense-making is not only accomplished by watching but by also with reference to background knowledge and inferences that build on these professionals’ understanding of human conduct.

LORENZA MONDADA (“Video Recording as the Reflexive Preservation and Configuration of Phenomenal Features for Analysis”) recommends what she calls a “praxeological approach” to video practices. On the basis of an ethnomethodologically inspired video analytic framework, she strives to take into account not only the question of how data are analysed, but also how they are produced. She addresses exactly what we referred to above as data sorts, i.e. the practices by which data are constructed. One kind of practice she refers to is the “praxeology of seeing”, i.e. the setting up of the video camera before the action, the kinds of camera movements and the filmer’s interaction with the camera. Moreover, she also hints at the fact that various professions work skilfully with video data, developing their own “professional vision”. In conclusion, she draws attention to the practices of editing video records.

HUBERT KNOBLAUCH points to the problems of analysing video data, proposing an approach he calls “videography”. His article explores the potential of combining ‘focussed’ ethnography with a microscopic analysis of video data. The programmatic title expresses the central importance of ethnographic field research for interpretive video-analysis. In combination with the attentive scrutiny of video sequences, ethnography is indispensable in order to make sense of and reconstruct the meaning of relevant details included in the recordings of social situations. Although video is an especially apt instrument for analyzing the details of action and interaction, a systematic collection of additional background knowledge is also of crucial importance. It is necessary to elucidate the visual aspects of the recordings, as the sequences are both situated and situative, that is both depending on and reflecting the larger social context.

In the final paper of the methodological first section, JÜRGEN RAAB and DIRK TÄNZLER suggest an approach they call “Video Hermeneutics”. This approach, based on Soeffner’s “structural hermeneutics”, has at its core a form of sequential analysis that attempts to reconstruct the range of readings, i.e. meanings, possible for single frames. By comparing different readings of key scenes, readings are excluded in order to arrive at a final, “objective” meaning. The interpretation is based on a “score” and proceeds by setting the context in parenthesis. They illustrate this approach in an analysis of two scenes of a television show.

Although they share the interest in methodology, the papers in the volume’s second section highlight the contribution of video analysis to specific *research fields*. Thus, DIRK VOM LEHN and CHRISTIAN HEATH (“Discovering Exhibits: Video-Based Studies of Interaction in Museums and Science Centres”) demonstrate how fruitfully video-analysis can be used for museum studies. The particular advantage of this method is that it allows us to study the conduct of visitors of museums arising with, at and around exhibits, in this way addressing the practice of aesthetics which has been

so often the subject of abstract theoretical debate. In order to do so, they analyse an example from a science museum. Moreover, their article also explains the reasons for conducting video analysis in general and some of the fundamental methodological issues relevant to such an analysis. Thus, they touch on the work of data collection as well as of transcription. They formulate three basic principles for data analysis: it is concerned with the indexical character of practical action, it considers social action as emergent and contingently accomplished, and it explicates the organisation through which participants produce particular actions.

CORNELIUS SCHUBERT (“Video-Analysis of Practice and the Practice of Video-Analysis”) also conceives of video analysis as addressing social practice. In particular, he turns to practices in medicine, that is to say in operating theatres in which actors are confronted with technology to such a degree that it seems plausible to him to frame technology as agents in order to clarify the practice observed. In his reconstruction of the practice of video-analysis, he stresses the role of Grounded Theory. Content logs resembling coding procedures may help the researcher to collect and compare data. Video also may be used as a medium for reflection since it allows for feedback and elicitation. Because video data are thus complemented by interview, observation and narratives, he proposes to call this method videographic video analysis.

ANSSI PERÄKYLÄ and JOHANNA RUUSUVUORI (“Facial Expression in an Assessment”) address a topic that had been prominent in psychology for a long time: non-verbal behaviour, or, in this case, facial expression. As opposed to the current attempts to analyse facial expression, they take an approach informed by conversation analysis. With respect to their data, which stems from “quasi-natural” conversations, they focus particularly on conversational assessments, i.e. the evaluations of persons and events that are described in conversational speech. In order to account for facial expressions found in the data, they develop a new transcription code which is added to the transcription of spoken utterances. Thus they demonstrate that the interpretation of facial expressions contribute significantly to assessments made within conversational contexts. Not only are facial displays coordinated interactively, but facial activities also incorporate the affective involvements of speakers with what is being assessed.

MONIKA WAGNER-WILLI bases her analysis of interaction in classrooms on the method of documentary interpretation suggested by Ralf Bohnsack (“On the Multi-dimensional Analysis of Video-Data. Documentary Interpretation of Interaction in Schools”). This method seeks to account for both the sequential aspects of video data as well as the simultaneity of visual information by distinguishing two dimensions of meaning: the explicit communicative dimension is at work when actors relate to the social role or the institutional order, whereas the conjunctive experiential space refers to the more implicit background commonalities of actors. She studies the threshold phase between breaks and lessons. This phase reveals itself as a transitional, liminal phase inbetween the conjunctive experiential space of the peer group and the communicative sociality of the school class.

BERNT SCHNETTLER (“Orchestrating Bullet Lists and Commentaries. A Video Performance Analysis of Computer Supported Presentations”) focuses on a relatively new option in face-to-face communication, which in many formally organized social situations quickly became something of an obligation, and that the author therefore claims to be a specific modern ritual: computer-supported presentations. Computer programs such as Microsoft’s PowerPoint offer speakers the opportunity to support their presentations with prefabricated and often animated visual impressions, i.e. diagrams and bullet lists, as well as more complex visual forms such as photos and video clips. Schnettler’s video performance analysis inquires into the specific new skills a speaker needs to coordinate different kinds of actions during his or her talk in order to gain social acceptance, and to prove him- or herself to be a competent performer. The case study of a computer-supported presentation arrives at the conclusion that ‘translating’ and ‘conducting attention’ are two core elements of a unique type of social action the author calls ‘orchestration’.

The contributions to the third part of the book share this interest in methodology, while at the same time drawing as well on a particular empirical field. In addition, they are characterised by their interest in the use of video for research (and the role of video for non-scientific practice). In studying classroom interaction, ELISABETH MOHN (“Permanent Work on Gazes. Video Ethnography as an Alternative Methodology”) calls for a manner of using video recordings which differs markedly from the “natural situation documentation” used by many. She draws on data collected while doing research in classrooms. Her argument is that the gaze, that is subjectivity of the video ethnographer as well as the visual character of these ethnographies, should be accounted for in the manner in which the data is collected and analysed. She proposes using video recordings as a form of field notes that follow the interests and the observational focus of the ethnographer. Thus, the camera moves according to what seems of importance to the ethnographer. As a result, the analysis, too, will be based on visual data, representing a departure from the word-centred report in favour of a visual display of the result.

In a similar vein, ERIC LAURIER and CHRIS PHILO (“Natural Problems of Naturalistic Video Data”) examine the question of the practical use of video in research settings. Although studying “a day in the life of the café”, they come to address what one used to call reactivity. Through their video recordings of people in cafés, they came to realise that the presence of the camera (and the absence of the ethnographer) is a constitutive feature of the setting recorded. Instead of getting rid of “reactivity” and thereby creating naturalness as the (artificial) absence of the recording device, they therefore turn to the ways in which subjects “react” to the presence of a video camera. The video, then, creates in their view a “videoactive context”, as Shrum, Duque and Brown (2005) would say. In fact, the subjects do not only react, the video triggers action on their part and thus contributes to the interaction.

Practice in a somewhat different vein is the topic of the paper by SIGRID SCHMID (“Video Analysis in Qualitative Market Research – from Viscous Reality to Catchy Footage”). She discusses the importance video has gained within the qualitative mar-

ket research. In this applied field, video analysis is employed in two ways: as a presentation tool and as an instrument of data collection. It is especially the ability to visualize consumers’ habits or a certain lifestyle within an everyday context that accounts for the advantage of video footage compared to more conventional research methods such as surveys, interviews or focus group studies. For the purpose of presentation, video footage can, she argues, first convey a holistic picture of the complex life worlds of individuals or groups of people. Hence, the production of these video images relies on typologies and findings generated in previous investigations. Secondly, and in a stricter sense, video is used in consumer studies as an instrument for data collection, but it also serves to identify new ideas for product innovation, the testing of prototypes or the further development of existing products.

The final paper returns to general issues of visual analysis. Restricting himself to photography, in his article on “Visual Sociology on the Basis of ‘Visual Concentration’”, HANS-GEORG SOEFFNER addresses an issue that is of major importance of any future video analysis. For, at this point in the field’s development, analysis is conducted by going out of one’s way via the use of written texts. Opposing this detour, SOEFFNER suggests that we could represent society through visual means themselves. Such an approach has been taken e.g. by Walker Evans and James Agee, by August Sanders in his famous portraits and also by Pierre Bourdieu. It is possible to use photography to the extent that one is successful in making explicit photography’s implicit catalogue of rules and its ‘interpretation of the world’ and thereby to methodologically control the visual displays. Even if SOEFFNER restricts his argument to photography, one could expect that it will become one of the touchstones in the field of video analysis due to its recommendations for finding ways of presenting findings in visual form and, probably, forms of “visual analysis”, as well.

Acknowledgments

We would like to express our gratitude to the *Deutsche Forschungsgemeinschaft* for the kind support of the Conference on Video Analysis as well as of our current research project. We also thank Eric Laurier and Polly Kienle for their help with the revision of some of the manuscripts.

References

- Aufschnaiter, S. v. & M. Welzel (ed.) 2001: *Nutzung von Videodaten zur Untersuchung von Lehr- und Lernprozessen*. Münster: Waxmann
- Banks, M. & H. Murphy, 1997: *Rethinking Visual Anthropology*. New Haven: Yale
- Bateson, G. 1958: Language and Psychotherapy: Frieda Fromm-Reichmann’s last project. *Psychiatry* 21, 21, 96-100

- Bateson, G. & M. Mead, 1942: *Balinese Character. A Photographic Analysis*. New York: New York Academy of Sciences
- Bergmann, J., T. Luckmann & H.-G. Soeffner, 1993: Erscheinungsformen von Charisma – Zwei Päpste. In: W. Gebhardt, A. Zingerle and M. N. Ebertz (ed.) *Charisma – Theorie, Religion, Politik*. Berlin/New York: 121-155
- Birdwhistell, R. L., 1952: *Introduction to Kinesics. An Annotation System for the Analysis of Body Motion and Gesture*. Louisville, Kentucky: University of Louisville
- Birdwhistell, R. L., 1970: *Kinesics and Context: Essay in Body-Motion Research*. Philadelphia: University of Pennsylvania Press
- Breckindrige, S. P. & E. Aboth 1910: Chicagoe's Housing Problems. Families in Furnished Rooms. *American Journal of Sociology*, XVI, 3: 289-308
- Büscher, M. 2005: Social life under the microscope? *Sociological Research Online*, 10, 1: <http://www.socresonline.org/10/1/buscher.html>
- Collier, J., 1967: *Visual Anthropology. Photography as a Research Method*. New York: Sage
- Collier, J., 1979: Visual Anthropology. In: J. C. Wagner (ed.) *Images of Information. Still Photography in the Social Sciences*. London: Sage, 271-282
- Collier, J. & M. Collier, 1986: *Visual Anthropology. Photography as a Research Method*. Albuquerque: University of New Mexico Press
- Curry, T. J. 1984: A Rational for Visual Sociology. *International Journal of Visual Sociology*, 1, 13-24
- Curry, T. J. & A. C. Clarke, 1978: *Introducing Visual Sociology*. Dubuque: Kendall/ Hunt
- Davies, C. A., 1999: Using visual media. In: C. A. Davies (ed.) *Reflexive Anthropology*. London, New York: Routledge, 117-135
- Denzin, N. K., 2000: Reading Film – Filme und Videos als sozialwissenschaftliches Erfahrungsmaterial. In: U. Flick, E. v. Kardoff and I. Steinke (ed.) *Qualitative Forschung. Ein Handbuch*. Reinbek bei Hamburg: Rowohlt, 416-428
- Dittmar, N., 2002: *Transkription. Ein Leitfaden mit Aufgaben für Studenten, Forscher und Laien*. Opladen: Leske + Budrich
- Ekman, P. & W. Friesen 1969: A Tool for the Analysis of Motion Picture Film or Videotapes. *American Psychologist*, 24, 3: 240-43
- Emmison, M. & P. Smith, 2000: *Researching the Visual*. London: Sage
- Erickson, F. & J. Schultz, 1982: The counsellor as gatekeeper. Social interaction in interviews. In: E. Hammel (ed.) *Language, Thought and Culture: Advances in the Study of Cognition*. New York: Academic Press, 237-260
- Finn, K. E., A. J. Sellen & S. B. Wilbur (ed.) 1997: *Video-Mediated Communication*. Mahwah: Erlbaum
- Fiske, J. 1998: Surveilling the City. Whiteness, the Black Man and Democratic Totalitarianism. *Theory, Culture and Society*, 15, 2: 67-88
- Fyfe, N. R. (ed.) 1999: *Images of the Street. Planning, Identity and Control in Public Space*. London, New York: Routledge
- Goffman, E., 1961: *Encounters (dt.: 1973)*. Indianapolis: Bobbs-Merrill
- Goffman, E., 1967: *Interaction Ritual. Essays on Face-to-Face Behavior*. New York
- Goffman, E., 1971: *Relations in Public*. New York: Basic Books
- Goodwin, C., 1981: *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press
- Goodwin, C. 1986: Gestures as a Resource for the Organization of Mutual Orientation. *Semiotica*, 62, 1/2: 29-49
- Gottdiener, M. 1979: Field research and video tape. *Sociological Inquiry*, 49, 4: 59-66

- Grimshaw, A. 1982: Sound-image data records for research on social interaction: some questions answered. *Sociological Methods and Research*, 11, 2: 121-144
- Heath, C., 1986: *Body Movement and Medical Interaction*. Cambridge: Cambridge University Press
- Heath, C., 1997: The Analysis of Activities in Face to Face Interaction Using Video. In: D. Silverman (ed.) *Qualitative Research. Theory, Method, and Practice*. London: Sage, 183-200
- Heath, C. & J. Hindmarsh, 2002: Analysing Interaction: Video, Ethnography and Situated Conduct. In: M. Tim (ed.) *Qualitative Research in Action*. London: Sage, 99-121
- Heath, C., H. Knoblauch & P. Luff 2000: Technology and social interaction: the emergence of workplace studies. *British Journal of Sociology*, 51, 2: 299-320
- Heider, K., 1976: *Ethnographic Film*. Austin: University of Texas Press
- Henney, L. M., 1986: *Theory and Practice of Visual Sociology (Current Sociology 34, 3)*. Lonon
- Holliday, R. 2000: We've been framed: visualizing methodology. *Sociological Review*, 503-521
- Jordan, B. & A. Henderson 1995: Interaction analysis: Foundations and Practice. *Journal of the Learning Sciences*, 4, 1: 39-103
- Knoblauch, H., 2003: *Qualitative Religionsforschung. Religionsethnographie in der eigenen Gesellschaft*. Paderborn, München, Wien, Zürich: Schöningh/UTB
- Knoblauch, H. 2004: Die Video-Interaktions-Analyse. *sozialer sinn*, 1, 123-128
- Koch, S. C. & J. Zumbach 2002: The Use of Video Analysis Software in Behavior Observation Research: Interaction Patterns of Task-oriented Small Groups. *FQS*, 3, 2
- Lacoste, M. 1997: Filmer pour analyser. L'importance de voir dans les micro-analysis du travail. *Champs Visuels*, 6, 10-17
- Lomax, H. & N. Casey 1998: Recording social life: reflexivity and video methodology. *Sociological Research Online*, 3, 2:
- Luckmann, T. & P. Gross, 1977: Analyse unmittelbarer Kommunikation und Interaktion als Zugang zum Problem der Entstehung sozialwissenschaftlicher Daten. In: H. U. e. a. Bielefeld (ed.) *Soziolinguistik und Empirie. Beiträge zu Problemen der Corpusgewinnung und -auswertung*. Wiesbaden: Athenaum, 198-207
- MacLean, A. M. 1903: The Sweat-Shop in Summer. *American Journal of Sociology*, IX, 3: 289-309
- Marks, D. 1995: Ethnographic Film: From Flaherty to Asch and after. *American Anthropologist*, 97, 2: 337-347
- Mead, M., 1975: Visual Anthropology in a Discipline of Words. In: P. Hockings (ed.) *Principles of Visual Anthropology*. The Hague, Paris: Mouton, 3-10
- Mittenecker, E., 1987: *Video in der Psychologie. Methoden und Anwendungsbeispiele in Forschung und Praxis*. Bern: Huber
- Pink, S. 2001: More visualising, more methodologies: on video, reflexivity and qualitative research. *Sociological Review*, 49, 1: 586-599
- Raab, J., 2001: Medialisierung, Bildästhetik, Vergemeinschaftung. Ansätze einer visuellen Soziologie am Beispiel von Amateurclubvideos. In: T. Knieper, Müller, Marion G. (ed.) *Kommunikation visuell. Das Bild als Forschungsgegenstand - Grundlagen und Perspektiven*. Köln: Halem, 37-63
- Raab, J. 2002: „Der schönste Tag des Lebens: und seine Überhöhung in einem eigenwilligen Medium. Videoanalyse und sozialwissenschaftliche Hermeneutik am Beispiel eines professionellen Hochzeitsvideofilms. *sozialer sinn*, 3, 469-495
- Rammert, W. & I. Schulz-Schaeffer, 2002: Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt. In: W. Rammert and I. Schulz-Schaeffer (ed.) *Können Maschinen denken? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt am Main: Campus, 11-64
- Ruby, J., 2000: *Picturing Culture: Explorations of Film and Anthropology*. Chicago: UCP

- Schefflen, A. E. 1965: The significance of posture in communication systems. *Psychiatry*, 27, 316-331
- Schnettler, B. 2001: Vision und Performanz. Zur soziolinguistischen Gattungsanalyse fokussierter ethnographischer Daten. *sozialer sinn. Zeitschrift für hermeneutische Sozialforschung*, 1, 143-163
- Shrum, W., R. Duque & T. Brown 2005: Digital video as research practice: Methodologies for the Millenium. *Journal of Research Practice*, 1, 1: Article M4
- Silverman, D. 2005: Instances or Sequences? Improving the State of the Art of Qualitative Research. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [Online Journal]*, 6, 3: Art. 30, available at <http://www.qualitative-research.net/fqs-texte/3-05/05-3-30-e.htm>
- Suchman, L. & R. H. Trigg, 1991: Understanding Practice: Video as a Medium for Reflection and Design. In: J. Greenbaum and M. Kyng (ed.) *Design at Work. Cooperative Design of Computer Systems*. Hillsdale: Lawrence Erlbaum, 65-89
- Thiel, T., 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. In: H. Keller (ed.) *Handbuch der Kleinkindforschung*. Bern, Göttingen, Toronto, Seattle: Hans Huber, 649-708
- Walker, N. 1915: Chicago Housing Conditions. Greeks and Italians in the Neighbourhood of Hull House. *American Journal of Sociology*, XXI, 3: 285-316
- Woodhead, H. 1904: The First German Municipal Exposition. *American Journal of Sociology*, IX, 4: 433-458

Methodologies of Video Analysis

Hubert Knoblauch

Videography

Focused Ethnography and Video Analysis

Introduction

Within the last two decades, video recording has become a more and more accepted instrument of data collection and analysis in the social sciences. However, while the collection of data using video appears largely unproblematic, the analysis of video data is more divided. This analytic variability can be initially simplified by distinguishing two types of analysis: the standardised (including automatised) and the interpretative. The standardised (and automatised) analytic approach begins with a number of pre-defined codes, constructed according to a given categorical frame that are “applied” to audiovisual recordings. After producing some kind of “intercoder reliability” or automatising the codes, audiovisually recorded interactions between pupils, for example, may then be coded as “supportive” or “non-supportive”, “aggressive” or “non-aggressive” (cf. Mittenecker 1987). In fields such as Computer Supported Collaborative Work or Human Computer Interaction, there are currently more than 40 software programs for standardised analysis available, most of them based on predefined categories (cf. Koch & Zumbach 2002). As useful as these studies may be for certain purposes, they differ from the requirements of the second approach, that of interpretive video analysis which is in the focus of this article (cf. Knoblauch, Schnettler and Raab, this volume). This is an analytic approach that starts from the assumption that actions are oriented to by meanings. Yet, while there are an impressive number of studies which have proven the necessity, importance and relevance of such interpretive video analysis, there have been a few attempts to delineate the methods of these studies, i.e. how to handle video for the purpose of social scientific research. Consequently, the field of interpretive, qualitative and naturalistic social scientific research still lacks any unified methodologies.

My aim in this paper is to build upon the disparate interpretative studies and propose a unified method. This will involve elucidating the methodological assumptions, and methodical steps, applied by a number of researchers in the field of interpretative video studies.¹ Video analysis in the social sciences, I will argue, can be considered as a

¹ I am fortunate to have had the chance to work with Thomas Luckmann in a series of projects involving video data, first as a junior researcher, later as a senior researcher. I am also indebted to John Gumperz whose research techniques I became familiar with at Berkeley. Most importantly, however, my gratitude goes to Christian Heath and his WIT group (especially Jon Hindmarsh, Paul Luff and Dirk vom Lehn) at King's College London whose dedication

kind of ethnography using video, to be more exact: a *videography*. As *videography*, the focus of video analysis is typically on what may be called visual conduct. Under the current heading of multi-modality, some researchers attempt to cut the forms of such visual conduct into slices, e.g. as gestures (Kendon 2004), kinesics (Birdwhistell 1970) or facial expressions (Ekman & Friesen 1969b). Whereas the attempts to re-synthesise the various slices or modes into a holistic picture, for example in analogy to the musical score, have proved to be utterly complex and so far, not really successful (cf. Luckmann this volume), approaches that take visual conduct as part of a social context seem to be able to cope much better with visual data. Based on Goffman's analysis of social situations and on conversational analysis, one of the central foci of interpretive, qualitative, social scientific video analysis has become interaction. By interaction, I refer to the fact that actors (e.g. those recorded) orientate their action at one another creating what Goffman (1981) has defined as a social situation. Interaction, however, is not restricted to people in co-presence. In fact, quite a number studies using video analysis (e.g. within the area of the Workplace Studies, cf. Heath, Knoblauch & Luff 2000) address forms of interactions in which technologies and visual representations of co-actors (such as documents, telephones, screens etc.) figure as "agents" (simulating co-presence). The focus on interaction, thus, also includes "interactivity" (Rammert & Schulz-Schaeffer 2002), i.e. forms of mediated interaction by means of technology, as well as the situative context. For this reason, it seems adequate to talk of *videography*.

In this paper, I will try to clarify in more detail what is meant by *videography*. Obviously, *videography* is only one of a number of methods applied when analysing video data. However, as I will argue in part 2, *videography* is probably the method that best exploits the potential of video for the social sciences since it combines (focused) ethnography with the "microscope of the social sciences". In a next part (3) I shall outline some methodological assumptions and methodical steps in order to make use of this microscope, particularly those addressing its sequential and visual features. And finally (4) I shall come back to the requirements on comparison and sampling that are, again, embedded in the ethnographic frame of *videography*.

Focused ethnography and video analysis

Studies in the social sciences that employ video data collection and its analysis frequently stress that they augment or complement their video recordings by field studies, participant observation or, more generally, ethnography. Indeed, if one looks more closely at video studies, and reflects upon one's own methodological practice, it becomes evident that the recording of video data is rarely developed in isolation. Most

to detailed video analysis has been so influential. I am also most grateful to Neil Jenkins for numerous valuable comments on the paper.

studies, in addition to video data, employ ethnographies which may include: participant observation, information from documents, comments from interviews and discussions, together with elicitations using the visual data recorded. Ethnography is of particular importance in those cases in which technological settings are studied; but also other areas in which the implicit and tacit knowledge of actors needs eliciting in order to understand the action at hand. For this reason, by the early 1980s Corsaro (1981) was proposing that there should be no video recording and analysis without undertaking prior ethnography – a procedure similar to what Albrecht (1985: 328f) already called "scouting". Indeed Erickson (1988) and Cicourel (1992) have stressed that this focusing process presupposes prior knowledge of and prior familiarity with the field and, therefore, prior ethnography. As Heath and Hindmarsh (Heath & Hindmarsh 2002: 107) emphasize, for the analysis of video recordings of naturally occurring activities "it is critical that the researcher undertakes more conventional fieldwork".

The bond between video analysis (in its social scientific, interpretive and naturalistic version) and ethnography has not occurred by chance, instead there is an intimate relation between video and ethnography (cf. also Shrum, Duque & Brown 2005). On the one hand, ethnography and video are both observational in a basic sense: audiovisual observation lies at the core of the activities of ethnography, and it is also audiovisual observation that is automatized or, to use Latour's (1986) terminology, "inscribed" into video-technology. There is, of course, a debate as to what degree this observation by means of video is "participatory", proactive, and contributes to the actions (Suchman & Trigg 1991, Jirotko & Goguen 1994) or "reactive", in affecting and manipulating the situation taped. However, while video equipment can be obtrusive and even obstructive to the local action, there are situations in which video recording may be less distorting than the presence of an (overt or covert) observer (see for example vom Lehn & Heath this volume).² Second, the inscription of the audiovisual allows one to address "natural situations" (see Knoblauch, Schnettler & Raab this volume) – the kind of situation that is also subject to ethnography. And finally, ethnography and video converge since they are both oriented to the conduct (or, as it was once more prosaically termed: the manners and folkways) of people in their ("natural") environment. As result of this convergence between video analysis and ethnography, I would argue, it is not misleading, but instead instructive, to talk of *videography* as the method to analyse people acting in social settings by video.

Videography admittedly differs to some extent from ethnography in a classical sense. For example, despite attempts to do multi-sited ethnographies of different organisational units by means of video, videographies typically do not aim at encompassing

2 As opposed to doing observations, the technologies of recording also relieve the researcher from other tasks and allow for ethnographical observations, questions and reflections while making the video records. Since the data collection is supported technically, researchers dispose of more time to observe specific features or to inquire into certain aspects of the already focused field.

large, locally distributed social structures, such as tribes, villages or cities. In fact, in contrast to such encompassing “conventional” ethnographies (as I will call them for the sake of brevity), videographies may be said to be *focused* in several ways.³

First, conventional ethnographies may be time intensive requiring a long, continual periods of fieldwork (as a rule for most students about a year). Whereas in contrast to this kind of experience-based ethnography, videographies tend to have shorter periods of data collection. Although even if fieldwork and data collection is over a shorter time period, the projects as a whole require similar amounts of time for completion.

Secondly, it has been suggested that the short periods of fieldwork for videographies can render them “superficial” or “quick and dirty” (Hughes, et al. 1994). This view, however, ignores the fact that, the short time periods covered are compensated for by another type of intensity: videography is enormously data intensive. This type of intensity refers not only to the huge amount of data collected in relatively short time, but to the fact that videography requires intensive and detailed data analysis to a degree unprecedented by conventional ethnographies based on field notes with written records.

Before clarifying the nature of video data analysis below, we should note that videography differs from most conventional ethnographies in a third sense; that of its scope or, to be more exact, focus. It lays a “focus on the particular”, i.e. the “particulars of situated performance as it occurs naturally in everyday social interaction” (cf. Erickson 1988: 1083). Videography, therefore, typically analyses structures and patterns of interaction, such as the coordination of work activities, the course of family arguments or professional meetings. Rather than studying, for example the police as a field, videography may turn to the question of how police officers do their patrols; or instead of studying youth clubs, it may focus on the question how young people perform at a certain event; or instead of studying the management of a company, it may focus on the meetings of managers. So whereas classical ethnography has turned towards social groups and social institutions, videographies are more concerned with specific actions, interactions and social situations.

It is this focus on actions, interactions and the social situation that motivates videographers to set their analyses within the framework of interpretive, qualitative and ‘naturalistic’ approaches in the social sciences. Based on the seminal work of scholars such as Mead and Goffman, particularly conversation analysis has influenced the analysis of video data (cf. Sacks 1992[1964ff]). However, due to the on-going restriction of “hard core” Conversation Analysis to audio data and its opposition to video data and background knowledge, the influence of conversation analysis on video analysis is less than it is with purely audio data (cf. ten Have 1999). However, as noted above, video analysis is not a unified approach and there are alternative ways to do video analysis (cf. the contributions to this volume by Mohn, Raab & Tänzler, Schnettler and Schubert

³ At this point I draw on my paper on focused ethnography, first published in German (Knoblauch 2001b) and in English (Knoblauch 2005).

etc.); yet we can say that videographies rely on ethnographic data to a varying degree – and that they account for the effect of ethnographic knowledge to varying degrees.

Video analysis: Sequentiality and Visuality

While ethnography may frame videography, its core lies in video-analysis. Video analysis builds on a number of features of the medium of video filming in general. On the one hand, the desire to record the “natural situation” is founded on the “mimesis” assumption that video recordings do indeed represent to some degree what is going on in situations that could be observed without a technological device. Despite the jeremiads on the “crisis of representation” and opposed to written documents of the situation, even the convinced post-modernist cannot deny that video recordings are accessible to other observers in ways that allow them to make new observations and interpretations and to give evidence for (and possibly against) other analyses. Compared to ‘plain’ observation, video recordings therefore appear more detailed, more complete and more accurate. Moreover, in a technical sense, they are a much more *reliable* kind of data than written fieldnotes since they allow analytic access for researchers who have not participated the data collection, i.e. independently of the person who collected the data (Peräklyä 1997).

Epistemologically, videography is not characterised by a suspension of belief in the existence of the things seen; on the contrary videography may be said to share a kind of “scientific realism” in that it assumes that people are existent and, that they have been conducting (acting) in ways that are open for reconstruction (capture) by video data.⁴ In this way, the video recording allows us to establish what Schutz (1962) has called subjective adequacy, i.e. a kind of correspondence between the statements of the researchers and those that are being researched and represented on video.

The advantage analysing video recordings is, that video is much more easily reproduced, manipulated and analysed than other visual data formats, e.g., film. The technical options of repeated viewings, dissection through slow motion and frame-by frame analysis, comparison through fast forward and data banks allow one to observe details that are not even visible to participants in such detail: often they cannot explain or even remember the bits and pieces of the visible behaviour that is accessible to the video analyst. One may anticipate that advances in digitalisation will enhance these options by enabling researchers to compare a number of different sequences simultaneously on one screen. Additionally, frames and sequences will be able to be linked and coded in a nonlinear ways and analysis will potentially become more and more visual (instead of verbal).

Finally, video analysis exploits another feature of this medium: its temporality. Like film, video is also defined by the temporal sequence of pictures (it is the fact that this

⁴ As to “scientific realism” and the distinction between these two levels cf. Luckmann (1978).

temporality is no longer built into the new digital storage medium that may cause the changes just mentioned above.) As a result of their temporality, pictures are watched in a consecutive *sequentiality*. It is this feature of sequentiality which has influenced the particular focus of much video analysis, i.e., actions and interactions, as the medium preserves the time structure of these temporal processes in a fashion unprecedented by earlier media (except of film). As video recordings mimic audiovisual conduct in time, they serve as a perfect medium to analyse the sequencing of action and the coordination of interaction through time. It is for this reason that video analysis can be regarded as akin to forms of analysis which are based on sequentiality, such as conversation analysis, objective or structural hermeneutics (cf. Soeffner this volume, Raab & Tänzler this volume). In other words, they consider sequentiality as the very structure by which social action, and thus social order is accomplished. Hence, video analysis starts with sequences of actions and interactions as its subject matter.

So we can briefly sketch how sequentiality can be considered the basis of interpretation and analysis. Firstly, interpretations of video recordings focus on audiovisual conduct. They assume that what is happening (and what is understood as happening) can be only understood if one looks closely at the actions, action sequences and interactions that are expressed in audiovisual conduct. Actions (as we call this basic category for the sake of convenience – without defining their boundaries) are assumed to be produced *methodically* in certain ways, and it is only by being performed in certain ways that certain things are brought about. Thus, a PowerPoint presentation (to take an example from our current research) would not be regarded as some background activity to what may be considered the core activity (e.g. “knowledge transfer”). Rather, the series of actions involved in doing the presentation will be considered as the very essence of this activity. The focus of analysis then lies in *how* those actions are being performed.

This is linked to an additional assumption which again lays stress on the audiovisual: Whatever is observable and understandable should not be considered as being due to external factors beyond the video recorded scene itself, such as “drives”, “subconscious desires”, attitudes or interests, but as motivated by the local sequence of action recorded. As Goodwin (1986) has shown, even alleged “adaptors” (as Ekman & Friesen 1969a call gestures such as scratching or coughing) turn out to be not just “outbreaks of nature”. Rather, they appear to be sensitively built into the sequential order of actions, and thus proving to be actions themselves. As opposed to conventional sociological lore, actions here are not seen as relating to other factors outside the specific situation in which they occur. This is at the heart of what we mean by sequential analysis where one considers any action (note: action not actors) as motivated by prior actions and motivating future actions. This assumption is often equated to the “situative” character of social action (Suchman 1987). However, I would consider situativeness a methodological principle by the analyst long as possible, and avoid treating it as a substantial feature of social actions. In fact, talking about the situative character precludes the observation of actions which cannot be interpreted by

situative, local resources only – and in which institutional regulations, social asymmetries and power may figure.

Reflexivity is another highly important methodological concept. By reflexivity we do not mean that actions are reflected consciously. On the contrary, most studies do address what may be called routinised, implicit knowledge or social practice. Reflexivity means that actors do not only act but also “indicate”, “frame” or “contextualise” how their action is to be understood and how they have interpreted a prior action to which they are responding.⁵ Thus we do not simply ask a question, we demonstrate that it is a question which we are formulating. It is because of this reflexivity that co-actors can understand what is meant by an action. By “investigating the methodological resources used by participants themselves in the production of social actions and activities” (Heath 1997: 184), it is also the reflexivity of the action that allows analytic interpretation.⁶

The possibility of interpreters and analysts making use of reflexivity does not only demand that they know the culture they are studying. It also demands that they understand situated action, rather than an a priori theory of communicative action. Such an understanding also means that analysts who have not participated in the phenomenon that was recorded are able to make sense of what is going on in the actions and interactions. This understanding has very practical aspects, for words and sentences have to be understood in order to be transcribed: one has to “see” the directions of gazes on the recordings or know what actors are referencing so that all essential parts (sentence, word, movement) of a sequence make sense for the participants. It is at this basic level of everyday understanding that ethnographic knowledge figures principally.

The necessity of interpretation proves that video analysis is basically a hermeneutic activity: the task set is not to only describe and explain “nonverbal” behaviour, but to (a) determine the knowledge which one needs in order to understand what is going on in a situation and (b) to identify the visible conduct that constitutes the situation.

Interpretation may be regarded therefore as the very first step of approaching the data, i.e. the tapes and transcripts. As interpretation that is related to the everyday understanding of what is seen and heard, it may be distinguished from analysis. By analysis we do not mean necessarily the distinction of different modes of audiovisual conduct (visual, verbal, paralinguistic etc.) – although it is necessary at certain points to focus on certain aspects, e.g. pointing gestures during PowerPoint presentations. Analysis, means identifying units of action and their interrelationships with one another in the sequence of their production. Whatever may appear to be a unit is to be interpreted in relation to what has been prior to it, and whatever the interpretation

5 This notion of reflexivity differs from notions of reflexivity that address the presentation of research (cf. Ruby 2000). Of course, reflexivity applies here, too, but this is not a feature of research only. What is meant by reflexivity is explained in more detail in Knoblauch (2001a).

6 In conversation analysis, it was assumed that reflexivity of spoken communication is framed in spoken communication only – an assumption that is not adopted by video analysis.

will be, it will be tested for its significance in the next unit (or turn). Hence interpretation is never only done retrospectively.

In conversation analysis, the basic unit of the sequence has been the “turn”. In fact, since video analysis still draws on transcripts of texts, turns at talk remain a starting point of the analysis. However, turns are not fixed action units but must be analyzed by looking at exactly how the utterances and the ways transitions and boundaries are produced. In order to understand the utterances, one may first try to look for acoustic cues, such as phrase intonations, pauses and rhythms.

Take as an example the following transcript⁷:

- 1 (3.3)
 2 M: ähm das is jetzt so die klassische Ausrüstung die einem zur Verfügung
ehm this is now the classical equipment which is available when
 3 *steht=wenn man eben Reis untersuchen möchte. Sie sehen hier auf*
one wants to investigate rice. You may see here on
 4 *der rechten Seite eben so ne ähm Lupe.*
the right hand side just such ehm a magnifying lens

The transcript shows a segment of text where M is speaking. By looking at the talk one realises that the description M provides is of some “equipment” he is referring to (2). By his address term (3) one can tell that he is addressing someone else and hints at the magnifying glass as part of the equipment with which to investigate rice. The sequence is structured so that he first gives a general title (“classical equipment”) to the topic (“investigate rice”); then he moves to singular item.

Scrutinising the transcript does not only allow us to become more familiar with the transcript and the spoken words (leading often to a continuous refinement and correction of the transcript). It is also the basis for analytical observations on the structure of the sequences studied. Moreover, the reading of the transcript serves a secondary function when watching the video and the visual conduct in more detail. The familiarity with the words spoken makes it easier to identify the sequence of events and “relate” the visual events with respect to the text. The video sequence can be repeatedly watched so as to discover the order of acoustic and visual events (what is happening when) and to identify when things are done or said.

It could be argued that this analysis could be accomplished without the transcript, but experience has shown that the availability of a written transcript functions as a kind of location device orientating the analyst in the video. The analysis of the consequent sequences thus moves from the (transcribed) written to the seen. However, the

7 An important requirement and resource of analysis consists in detailed transcriptions of data. Transcriptions (which typically involve a lot of work and confront the researcher in a very intensive way with the data) include paralinguistic and prosodic features (Cook 1990) which, however, do not figure prominently in the example given. As the transcription of data is as time-extensive as is data analysis, it seems decisive to develop strategies of sampling and selecting relevant data.

visual is not treated merely as an addendum to the spoken. Rather, in watching the video recordings, one frequently discovers additional sequences which allow one to make sense of prior interpretations or sequences.

Without the video data and without ethnographic knowledge it would be difficult to tell that M is not just talking; but that in addition he uses a PowerPoint presentation on which what he is talking about is represented.⁸ Moreover, that he is pointing at the screen by means of a laser pointer. Thus, the deictic noun “das ist” (“that is”) is part of introducing the new slide whereas with the deictic

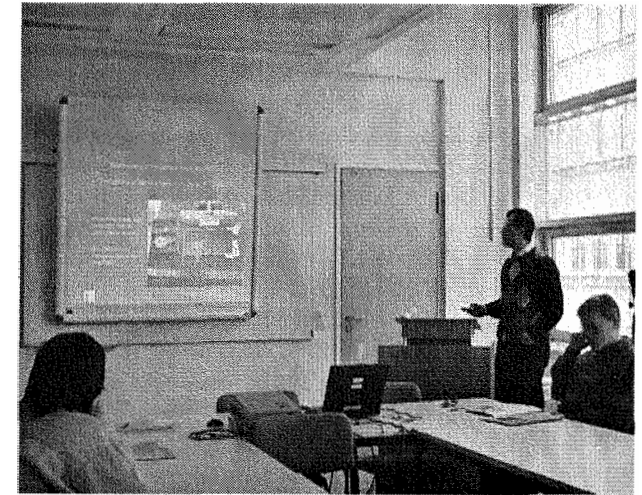


Image 1: PowerPoint presentation with pointer

noun “Sie sehen hier” (“you may see here”), M points at the items he is going to talk about.

In looking at both, the spoken and the visual conduct, one can then discern that the visual conduct is constructed in such a way as to support the text the speaker is producing, and vice versa. The PowerPoint slide shown on the screen (Image 1.) is, in fact, quite complex, but the ways of highlighting the text as well as pointing to the screen create an object which has not been there before and which then may serve as a point of reference in the next turn – the turn itself being constructed by visual (body movement, new slide) and verbal (pause) features. Analyses of this kind aim at the sequential structure of events, the construction of meaningful units, participation structures, the spatial organisation of activities, and the role of artefacts for performing activities etc. (Jordan & Henderson 1995).

Thus, the validation procedure for backing-up the analysis draws on what we may call an immanent criterion of sequentiality. Analysis depends on what occurs audiovisually: every utterance, every gaze, every move of the body or the head is being taken

8 The data pertains to a research project supported by the Deutsche Forschungsgemeinschaft (DFG) called “Die Performanz visuell unterstützter mündlicher Präsentationen. Gattungsanalytische Untersuchung einer paradigmatischen Kommunikationsform in der ‘Wissensgesellschaft’” (“The Performance of visually supported oral presentations: Genre analytical investigation of a paradigmatic form of communication in the ‘knowledge society’”).

into consideration if, and how, it forms part of a sequence. Sequences are not identified independently of the sequences prior to, or subsequent to themselves, but by their immediate local context. One needs not to consider everything on the screen, but only that which stands in a recognisable relation to what came before and what comes after. Thus, turns at talk may not be the relevant units of sequences at all. But whatever may be considered as unit must be shown to be bounded by audiovisual conduct itself.

This criterion of sequentiality is complemented by a second criterion: what is of importance in the visual must not be speculated about but be indicated by the actors themselves. Do actors look at the screen before they push a button? Does it ring before A walks away? Does A look at B before A moves towards C? Schegloff (1992) called this the criterion of relevance, i.e. what is of relevance for the analyst must be shown to be of relevance for the actors. Or, as Goodwin (2000: 1508f) formulates: "Rather than wandering onto field-sites as disinterested observers, attempting the impossible task of trying to catalogue everything in the setting, we can use the visible orientation of the participants as a spotlight to show us just those features of context that we have to come to terms with if we are to adequately describe the organization of their action". This task can be accomplished by an analyst who knows both the field and the video data well. In addition, it is beneficial to regularly hold data sessions in which the analytical observations of sequential orders are subjected to the critical gazes of other observers who can gradually become familiar with the data. Finally, data workshops with other researchers and students who are less familiar with the data can be organized which may provide additional perspectives and/or help in testing more encompassing observations.

Up to this point, the analysis has focused only on the sequence of actions (including technical devices, such as the screen on which the items are represented). This focus must not be considered a problem since the example refers to a focused kind of interaction in which participants themselves establish a common focus. Things change however if the events recorded are not focused interactions, if spoken words are only of minor importance and if, therefore, the visual gains in importance. For the spoken may be represented in a sequential and linear order corresponding to time, that is a diachronic order. Yet the visual denotes an additional synchronous dimension of simultaneity, i.e. the visually represented courses and resources of action, features of actors and thus the contexts may not be organised on a turn-by-turn basis.

One way to approach this problem in video analysis has been suggested by Goodwin (2000). He suggests that semiotics may be able to grasp these visual features. Thus, talk is embedded in multiple sign systems (such as graphic codes, gestures and other features of the environment). To Goodwin, actors orient to what he calls "semiotic fields" that include different kinds of sign phenomena instantiated in diverse media. In accordance with the principle of relevance, these semiotic fields may be of local relevance in that the actors demonstrably orient towards them (Goodwin calls this "contextual configuration").

There is no doubt that the concept of semiotic field works in cases of well developed sign systems (such as graphic systems of professional experts or the sign elements of

children playing hopscotch). One must, however, question if semiotics will help to remedy the problem of visibility in general, since semiotics presupposes that the signs visible are organised in a more or less systematic way. Even if one may admit that words form part of a system, one may doubt that gestures can be considered as forming a real system, let alone other visual elements (such as expressions, clothing, furniture, the order of things in space etc., cf. Hodge & Kress 1988).

Returning to Ethnography

Validation by the criterion of relevance may often turn out to be an intricate task. In order to clarify the meaning and signification of visual elements of the recorded data, therefore, one further possible procedure is elicitation, auto-confrontation or video-based interviewing.⁹ This means that the actors involved in the recorded data are presented with the recordings themselves. Thus, Schubert for example (this volume) presented medical personal he had been videoing with recordings of the their operating theatre work and asked them to clarify actions they were involved in. This method does not only potentially allow the reconstruction of actors' perception and orientation in the captured action. It may be seen to also give access to background knowledge relevant to understanding what is going on as actors may be able to explicate the functioning and significance of items visible in the scene, but not available to the analyst.

The importance of background knowledge that elucidates visual aspects of the recordings proves again the importance of ethnography for doing video analysis. For it is by way of observation, interviews, expert interviews etc., i.e. ethnography, that we get familiar with, and make sense of, the settings in which we produce the video recordings. We acquire the knowledge necessary to understand the audio-visual action temporally, i.e. the series of actions (which, in many empirical cases, also involves intricate technologies and their activities). Thus videography still addresses the emic perspective of the natives' point of view, yet in a very specific sense: specified with respect to certain situations (situated), activities and actions. This does not mean that one needs to reconstruct the stock of cultural knowledge (i.e., members' knowledge) necessary to act in the domain as a whole. The task of the researcher is to acquire sufficient knowledge, particularly those elements of knowledge, partly embodied, relevant to the activity on which the study focuses. When studying technological activities, for example, especially those elements of knowledge necessary and relevant to understand the practices involved in handling that technology are sought.

9 Cf. Bayart, Borzeix & Lacoste (1997). The method goes of course back to Jean Rouch, cf. Jackson (2004).

Typically, it is these types of sequences that are the focus of videographies: Ways of managing technologies, forms of cooperation between personal and their use of technology at certain circumstances (for example in routine actions or crises interventions), kinds of interactions between lay and professional actors etc. The focus may vary strongly depending on the research topic and the structure of the field. Nevertheless, one of the principal goals of the analysis will be to identify common features of what is being studied. In order to do so, the analysis will not only need to interpret and analyse single cases, but will also utilize case comparisons. Such comparison means that similar cases, according to the features that are identified in the analysis, can be located. Comparisons can be associative and look for similarities, or they will be identifying minimal and maximal contrasts. In general, these comparisons will help to determine certain patterns in the sequences studied, be they institutionalized (as e.g. certain organisational "problem solutions") or context dependent and situational.¹⁰ These sequences may consist of series' of individual actions, of interactions between various actors and of actions of technologies. Therefore, analysis tends to demonstrate the kinds of interrelationship of these actions in the situations recorded.

For this reason, single instances may be compared which are stored in the data corpus. In order to retrieve these instances, one draws on a content log.¹¹ On the basis of such a content log, activities can be analysed across a given corpus of video data, e.g. locating forms of pointing from a corpus of data of video recorded PowerPoint presentations in order to identify the specific features of pointing with this technology, the role of the technology and the effect it has on presentations.

As is often stressed, these situations do have *situative*, contingent properties realized only in the situations which provide a resource for the actors. On the other hand, one should be aware that situations also include *situated* properties, that is, they have properties that are observable across situations and form part of "larger context" (Goffman 1983), be it the kind of rooms and their micro-ecologies; the technologies available, the ranks of persons present and their representations as well as features of the decorum; the situations recorded form part of settings, institutions, organizations and other contexts. There is no doubt that these contexts are subject to the ethnography, be it social welfare agencies, underground stations or management offices. Thus, instead of just comparing interactive sequences, the analysis may then return to the larger context which has been subject to the ethnography (Whereas the first comparison, then, may be regarded as analysis within a data sort, this one resembles to a triangula-

10 As far as we look for patterned features, we investigate what we call "communicative patterns", that is forms of interaction that exhibit communal structures beyond the situative actions that relate to extra-situational functions and social structures. Cf. Knoblauch & Guenther (1995).

11 A content log contains the temporal sequence of events, a rough transcription of activities, gestures and talk, reflections and codings of sequences according to the research topic. Cf. Jordan & Henderson (1995).

tion across data sorts, cf. Flick 2004: 36ff). This context may be kept constant in some studies. Thus, one may concentrate on counseling activities in social welfare agencies, on cooperation between actors in underground control rooms or in ticket-sale interactions between service personal and clients in stations. In these cases, the ethnographic details which enter into the background knowledge of the analysts are to some degree constant. In other cases, however, the contexts in which certain activities that are shown or claimed to be recurrent may vary across various settings, institutions or milieus. To give an example, we may ask if PowerPoint presentations in academic settings differ from the same genre in private companies or in administrative settings. In order to avoid misunderstandings these different settings are not considered as "external factors", but they may prove to denote relevant settings according to which certain interactional sequences may differ or (what could be of even more interest) exhibit strong similarities.¹² Analysis, then, means not only considering interactive situations but also the larger social context (and showing how the former constitute the latter). It is for this reason that analysis can be seen to begin in the data collection process and to guide the sampling of video recordings at every stage. Throughout the research process it is related to, and dependant upon, ethnography. It is due to this close relationship between video analysis and ethnography that I suggest the term videography adequately accounts and describes the unified core of what constitutes the interpretative video analysis methodology.

References

- Albrecht, G. L. 1985: Videotape Safaris: Entering the Field with the Camera. *Qualitative Sociology*, 8, 4: 325-344
- Bayart, D., A. Borzeix and M. Lacoste 1997: Les traversées de la gare: Filmer des activités itinérantes. *Champs visuels*, 6, 75-90
- Birdwhistell, R. L., 1970: *Kinesics and Context: Essay in Body-Motion Research*. Philadelphia: University of Pennsylvania Press
- Cicourel, A. V., 1992: The interpretation of communicative contexts: Examples from medical encounters. In: A. Duranti and C. Goodwin (ed.) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 291-310
- Corsaro, W. A. 1981: Something old and something new. The importance of prior ethnography in the collection and analysis of audiovisual data. *Sociological Methods and Research*, 11, 2: 145-166
- Ekman, P. and W. Friesen 1969a: The Repertoire of Nonverbal Behavior: Categories, Origins, Usage and Coding. *Semiotica*, 1, 63-68
- 12 It is the goal of genre analysis to identify these kinds of similarities and differences and show in which way they contribute to the construction of situations and larger social structures. Cf. Knoblauch & Günthner (1995).

- Ekman, P. and W. Friesen 1969b: A Tool for the Analysis of Motion Picture Film or Videotapes. *American Psychologist*, 24, 3: 240-43
- Erickson, F., 1988: Ethnographic description. In: U. Ammon (ed.) *Sociolinguistics. An International Handbook of the Science of Language and Society*. Berlin, New York: de Gruyter, 1081-1095
- Flick, U., 2004: *Triangulation. Eine Einführung*. Wiesbaden:
- Goffman, E., 1981: *Forms of Talk*. Philadelphia: University of Pennsylvania Press
- Goffman, E. 1983: The Interaction Order. *American Sociological Review*, 48, 1-17
- Goodwin, C. 1986: Gestures as a Resource for the Organization of Mutual Orientation. *Semiotica*, 62, 1/2: 29-49
- Goodwin, C. 2000: Action and embodiment within situated human interaction. *Journal of Pragmatics*, 32, 1489-1522
- Günthner, S. and H. Knoblauch 1995: Culturally Patterned Speaking Practices - The Analysis of Communicative Genres. *Pragmatics*, 5, 1: 1-32
- Heath, C., 1997: The Analysis of Activities in Face to Face Interaction Using Video. In: D. Silverman (ed.) *Qualitative Research. Theory, Method, and Practice*. London: Sage, 183-200
- Heath, C. and J. Hindmarsh, 2002: Analysing Interaction: Video, Ethnography and Situated Conduct. In: M. Tim (ed.) *Qualitative Research in Action*. London: Sage, 99-121
- Heath, C., H. Knoblauch and P. Luff 2000: Technology and social interaction: the emergence of 'workplace studies'. *British Journal of Sociology*, 51, 2: 299-320
- Hodge, R. and G. Kress, 1988: *Social Semiotics*. Ithaca and New York
- Hughes, J. A., V. King, T. Rodden and H. Anderson, 1994: Moving out of the Control Room: Ethnography in System Design. In: R. Futura and C. Neuwirth (ed.) *Transcending Boundaries. Proceedings of the Conference on Computer Supported Cooperative Work*. Chapel Hill: 429-439
- Jackson, J. 2004: An ethnographic flimflam: Giving gifts, doing research, and videotaping the native subject/ object. *American Anthropologist*, 106, 1: 32-42
- Jirota, M. and J. Goguen (eds.) 1994: *Requirements Engineering: Social and Technical Issues*. London
- Jordan, B. and A. Henderson 1995: Interaction analysis: Foundations and Practice. *Journal of the Learning Sciences*, 4, 1: 39-103
- Kendon, A., 2004: *Gesture. Visible Action as Utterance*. Cambridge: Cambridge University Press
- Knoblauch, H., 2001a: Communication, contexts and culture. A communicative constructivist approach to intercultural communication. In: A. di Luzio, S. Günthner and F. Orletti (ed.) *Culture in Communication. Analyses of Intercultural Situations*. Amsterdam, Philadelphia: John Benjamins, 3-33
- Knoblauch, H. 2001b: Fokussierte Ethnographie. *Sozialer Sinn*, 1: 123-141
- Knoblauch, H. 2005: Focused Ethnography. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [Online Journal]*, 6, 3: Art 44, Available at: <http://www.qualitative-research.net/fqs-texte/3-05/05-3-44-e.htm>
- Knoblauch, H., B. Schnettler and J. Raab, this volume: Video-Analysis. Methodological Aspects of Interpretive Audiovisual Analysis in Social Research. In: (ed.)
- Koch, S. C. and J. Zumbach 2002: The Use of Video Analysis Software in Behavior Observation Research: Interaction Patterns of Task-oriented Small Groups. *FQS*, 3, 2
- Latour, B., 1986: Visualization and Cognition. Thinking with eyes and hands. In: H. Kucklick and E. Long (ed.) *Knowledge and Society. Studies in the Sociology of Cultural Past and Present*. New York: Jai, 1-40
- Luckmann, T., 1978: Philosophy, Science and Everyday Life. In: T. Luckmann (ed.) *Phenomenology and Sociology*. Harmondsworth: Penguin, 217-253

- Luckmann, T., this volume: Some remarks on scores in multimodal sequential analysis. In: H. Knoblauch, B. Schnettler, H.-G. Soeffner and J. Raab (ed.) *Video-analysis*.
- Mittenecker, E., 1987: *Video in der Psychologie. Methoden und Anwendungsbeispiele in Forschung und Praxis*. Bern: Huber
- Peräkylä, A., 1997: Reliability and validity in research based on tapes and transcripts. In: D. S. Silverman (ed.) *Qualitative Research*. London: 199-220
- Raab, J. and D. Tänzler, this volume: Video-Hermeneutics
- Rammert, W. and I. Schulz-Schaeffer, 2002: Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt. In: W. Rammert and I. Schulz-Schaeffer (ed.) *Können Maschinen denken? Soziologische Beiträge zum Verhältnis von Mensch und Technik*. Frankfurt am Main: Campus, 11-64
- Ruby, J., 2000: *Picturing Culture: Explorations of Film and Anthropology*. Chicago: UCP
- Sacks, H., 1992[1964ff]: *Lectures on Conversation. Edited by Gail Jefferson and Emanuel A. Schegloff*. Oxford: Blackwell
- Schegloff, E., 1992: On talk and its institutional occasions. In: P. Drew and J. Heritage (ed.) *Talk at Work. Interaction in Institutional Settings*. Cambridge: 101-136
- Schubert, C. this volume: Video-Analysis of Practice and the Practice of Video-Analysis.
- Schutz, A., 1962: Common Sense and Scientific Interpretation of Human Action. In: (ed.) *Collected Papers I: The Problem of Social Reality (ed. by Maurice Natanson)*. The Hague: Nijhoff, 3-47
- Shrum, W., R. Duque and T. Brown 2005: Digital video as research practice: Methodologies for the Millennium. *Journal of Research Practice*, 1, 1: Article M4
- Soeffner, H.-G., this volume: Visual Sociology on the Base of 'Visual Concentration'
- Suchman, L., 1987: *Plans and Situated Actions. The Problem of Human-Machine Communication*. Cambridge
- Suchman, L. and R. H. Trigg, 1991: Understanding Practice: Video as a Medium for Reflection and Design. In: J. Greenbaum and M. Kyng (ed.) *Design at Work. Cooperative Design of Computer Systems*. Hillsdale: Lawrence Erlbaum, 65-89
- ten Have, P., 1999: *Doing Conversation Analysis. A Practical Guide*. London
- vom Lehn, D. and C. Heath, this volume: Discovering Exhibits: Video-Based Field Studies in Museums and Science Centres

Video Hermeneutics

Video technology offers new perspectives for sociologists. Above all ethnographic approaches, which have been eager to adapt the new technique as a means of data production, have had a crucial impact on the field of qualitative social research. *Video hermeneutics* have recently been developed as a procedure for the understanding of audiovisual data. Sociological hermeneutics are closely related to the theory of the social construction of reality (Berger & Luckmann 1966). The fundamental idea of this approach is to consider social data as manifestations of the protagonists' perception and recognition of reality as well as of their self-representation and self-interpretation. Consequently, *video hermeneutics* as a reconstruction procedure shows how facts are fabricated by human beings under certain socio-historical conditions. It also obliges the researcher to take on a self-reflexive stance and take into account his or her subjective presuppositions under which he himself or she herself constitutes the reality he or she is observing. Only in this way social scientists do justice to the intention of an ethnographic approach.

In the following, we outline some characteristic features of the methodology (I) and the method (II) of *video hermeneutics*. A few, but significant theoretical elements of the hermeneutical procedure are presented. Problems of the transcription, the description, and the interpretation of audiovisual data are discussed more or less theoretically. Finally, we demonstrate the practice of *video hermeneutics* in a brief case study (III).

Understanding Images as a Way of Understanding Culture

Modern societies are characterised by the increase of medial and visualised forms of communication which have deep impact on social relations. Medial representations of reality tend to overlay the 'natural' perceptions generated by the human senses. In other words, media products not only increasingly surround people in their everyday-life, but photographs, movies, TV-broadcasts, video-productions, and virtual computer-worlds influence their perception of reality fundamentally. Half a century ago Marshall McLuhan formulated his favourite thesis on communication that "the medium is the message" (McLuhan 1964), and he recently contributed the prophetic vision of a "global village" (McLuhan & Powers 1989). But it is still a challenge for sociologists to explore the ramifications of audiovisual media to culture and society.

Researchers must answer questions like: To what extent do technical constructions of reality alter the *forms* of human self-interpretation and self-representation? How do the audiovisual media shift and extend the *potential* for the human construction and attribution of meaning? And not least, which new requirements for the interpretation

and which new challenges to the understanding of meaning come into being in everyday-life (e.g. in the reception of mass media), as well as in social-scientific interpretation?

In the humanities and the social sciences, images were overshadowed by a focus on analysis of written language and of texts for many years (Goody 1981, 2000). However, since the 1990's, boundless discussions on the *pictorial turn* (Mitchell 1994), the *imagic turn* (Fellmann 1995), and the *iconic turn* (Boehm 1994) have indicated that the paradigm of the *linguistic turn* (Rorty 1967) and therefore of speech and text are beginning to lose their privileged positions. All these turns suggest a change of paradigm in the humanities and the social sciences and stand for the rediscovery of the relevance of images – especially and above all in view of the new forms and the new dimensions of pictorial communication and audiovisual recording techniques available in modern societies. They are propelled by the attempt to at last do theoretical, methodological, and methodological justice to the historical, cultural, and social function of images.

In his famous definition, Max Weber calls for “a sociology [...] as a science concerning itself with the interpretative understanding of social action and thereby with a causal explanation of its course and consequences” (Weber 1978: 4). *Video hermeneutics* as a new branch of sociological hermeneutics represent perhaps one of the most thorough and probing attempts to realize Max Weber's program of an interpretative sociology (Soeffner 2004). Furthermore, for *video hermeneutics*, human perception and social action, the human constitution of meaning, and the social transmission of knowledge are always conducted through symbolic forms. In brief, social action takes place by means of symbolic exchange (Mead 1967, Schutz & Luckmann 1974, 1989, Berger & Luckmann 1966). However, even hermeneutic approaches in the social sciences fail to consider sufficiently or simply oversee the fact that images are a quite particular medium for the constitution of meaning and for the social construction of reality. Most importantly, images call for entirely different approaches to the process of interpretation than those used for speech and texts.

Developing Ernst Cassirer's idea of symbolic forms, Susanne K. Langer makes the fundamental and therefore useful differentiation between *discursive* and *presentative symbolism* (Langer 1967). In language, only those things can be expressed that fit into a discursive order. A discursive order is a linear sequence of already significant elements. Any idea that cannot be adapted to this form of projection is unutterable and cannot be communicated by means of words and sentences. In contrast, images are characterised as a simultaneous and integral symbolism that Langer calls presentative. All elements that constitute the meaning of an image are simultaneously present.

Applying Langer's differentiation to the phenomenon of video, we can say that audiovisual data represent a *hybrid* of discursive and presentative symbolism. In this kind of data simultaneity and succession of symbols are affiliated and constitute a specific mode of signification on three levels. Firstly, these are *events and objects before the camera* – always accompanied by language (spoken or written, dialogues or monologues, narratives or commentaries), sounds, and music. Secondly, *camera actions* are

present – the image itself is moving (changing camera positions, pan shots, moving cameras, and zooms). Thirdly, *editing techniques* mold the data – offering an opportunity to separate and to link units of meaning (cutting and montage) as well as to embellish, to pad, and to optimise the material in many ways.

This brief phenomenological description has crucial consequences for the methodology of *video hermeneutics*. The integration and the combination of different levels of action and different symbolic forms make audiovisual recordings the most complex procedure of data-production in the social sciences. However, it is for this reason that audiovisual recordings present the greatest challenge for interpretation. Due to the richness of their content, to their complexity, and to their variability audiovisual data require appropriate procedures for transcription, description, and interpretation both in their discursive succession and in their presentative simultaneity. In the following, we give a rough sketch of the three central methodic principles of *video hermeneutics*: sequentiality, parenthesis of context, and contrasting.

Methodic Principles

Sequentiality

At the core of *video hermeneutics* lies the *sequential analysis*. The meaning of sequentiality here is twofold. On the macro-level it means, that not the whole datum – for example the film or video-record as a totality – but only key scenes are interpreted. On the micro-level it means, that these key scenes are interpreted step-by-step, i.e. picture-by-picture.

The key scenes are selected by an interpretation group – the procedure functions optimal only when a research group conducts the analysis, not so well by a lonesome researcher – after they have watched the recorded data in its entirety. On the macro-level there are really no guidelines that determine the choice of a particular sequence for detailed analysis. Instead, the researchers themselves decide spontaneously in view of their project goals and thematic focus, which sequence they find most intriguing and prefer to interpret. The idea of the sequential analysis is, to conceive the motives underlying the intuitive selection of a key scene by the interpreters. The understanding of the interpreter's attraction to the key scene coincides with the attempt to reconstruct and explain the structure of the datum. The restriction to an analysis of key scenes is a consequence of the aim of providing a non-reductive and extensive interpretation of the data in all recognisable details and aspects (events and objects before the camera, camera actions, and editing techniques).

Assuming that social reality is a meaningful and regulated order, at the micro-level we follow the temporal sequence of the data in an exact and strict manner. To this end, the researchers 'freeze' the flow of pictures and create motionless stills. At this point, they begin to describe the stills in all details. The descriptions are noted in a

formula developed for the transcription of film and video-documents called a *score* (see the score in the appendix, also Bergmann et al. 1993 and Luckmann in this volume).

In general, the *score* is a technical instrument used to translate visual and acoustic data into written language. For hermeneutics, *text* is a fundamental prerequisite to interpretation. The score requires that the interpreters not only verbalize their thoughts in spoken language, but that they also risk a first translation of their observations into written text. The schematic structure of the score supports them in this effort through the analytic subdivisions it requires. In the first steps of the analysis, they do not yet have to represent the complex relationships within the sequence as a whole. Instead, they can limit themselves to selecting single elements of the action and translating them into written text as a first approximation of the meaning of the sequence. In short, the score is the first *locus of translation*, but it accomplishes this purpose in a special manner adequate to the nature of the video-material. Therefore, the construction of the score must be flexible, i.e. it must adapt to historical changes in general and to technical developments especially in the audiovisual media itself, as well as to the research interests and to the formulation of the scientific question.

Due to the fundamental hybrid character of audiovisual data, the score must record all actions in their simultaneous and successive order. To realise this methodic principle all events are distinguished in their audiovisual dimensions, described in detail and noted in different columns in accordance with the research interests (i.e. setting, camera, body, head, speech, and music). In some cases, it seems necessary to distinguish between different kinds of bodily-actions, in other cases not. The linear *succession* of the events in the dimension of time is fixed vertically ('top-down') in the spatial dimension of the columns, which run parallel to each other. In the horizontal dimension, therefore, all events are shown in the *simultaneity* of their appearance. It is possible then to reconstruct, for example, what happens with the body or on the face of an actor while he is speaking. Thus, the interpreters are forced to look more closely at the data than is perhaps usual, and the score represents an instrument with which they can check for redundancies in the recorded data of the event at any point.

Lastly, the score is an elementary means for monitoring one's own interpretation and thus an analytical tool for a methodic approach as stipulated by Grounded Theory (Glaser & Strauss 1967). It is also an instrument for third persons to *check* and to *confirm* the single steps, as well as the entire interpretation, according to general methodological standards.

In principle, *video hermeneutics* perceive the practice of notation and that of interpretation as inextricably linked to one another. A notation is not only a description, but at the same time the beginning of an interpretation of the observed social action.

The score is an intermediate step enabling the interpreters to progress from a highly complex, synaesthetic perception of social action, to finally arrive via a structured analysis at a description and interpretation in the form of a coherent text. It is obvious that the score can replace neither the real object of the analysis nor the interpretation

in the form of a text. The last reference for all understanding and interpretation is the datum in its original appearance and not a text of any kind. But, nevertheless for hermeneutics text is a fundamental condition for interpretation.

In contrast to understanding as a spontaneous act, interpretation is an experimental setting in which hypotheses about the observed social reality are constructed on the basis of a text. We can say that the sensual phenomena of the audiovisual datum are translated in the score as an abstract form of text. In the sense of Alfred Schutz (1953) a 'natural' "first order construction" is transformed into a "second order construction". Two aims are realised in this way. The data are interpreted in accordance with the research interest, and the interpretation is set down as an adequate representation of the datum in its 'natural' order. Thus, methodic constructivism is linked with epistemological realism.

Parenthesis of Context

Establishing *parenthesis of context* is perhaps the most important procedure in sociological hermeneutics. However, it is also one that is most likely to be misunderstood. Beginning with an image-by-image interpretation, we do not attempt to secure and install *one* particular interpretation of the social action. On the contrary, we intend to delay the process of attributing meaning to our data, and remain in the role of inquirers as long as possible.

Sociological hermeneutics are not procedures for the empirical testing of theoretical hypotheses. Instead, they serve to generate conjectures on the subject. By *parenthesis of context* we free ourselves from common sense and all standardised explanations in everyday-life and scientific understanding – i.e. from orthodox and dogmatic world-views – to achieve a fresh perspective offering new and innovative insights into social reality. In short, *parenthesis of context* is the methodological technique that enables us to become conscious of and to account for the social construction of contexts, in which the recorded data gain meaning and relevance. Just as was crucial to Goffman's sociological approach of the understanding of human action (Goffman 1974), we ask ourselves: What is really going on here? Why are the protagonists behaving the way they do, and why not otherwise? And: why have they chosen these options while rejecting others?

In order to find answers to these questions, it is important to design possible contexts in which certain data can occur and fit best. We are obliged to develop a multitude of interpretations that may indeed be mutually contradictory. However, in order to do so, we must take recourse to an artifice that, as Karl Mannheim put it, dissociates us from the principal 'position-boundedness' (*Standortgebundenheit*) of our own knowledge (Mannheim 1960). Through this artificial installation the researcher is alienated from the 'natural' given data as well as from the ordinary worldview he takes for granted. He is thus forced to create a multitude of possible interpretations of the data under changing fictitious contexts as in an experiment 'in vitro'.

Seen this way, *parenthesis of context* means that, in the first step of the interpretation, we ignore as much of our knowledge about the data that constitute the data we are

studying as possible. We thereby achieve distance to what had seemed self-explanatory, above and beyond the image-by-image interpretation already mentioned.

According to Hans-Georg Gadamer, “understanding begins [...] when something addresses us. This is the primary hermeneutical condition. We know now what this requires, namely the fundamental suspension of our own prejudices. But all suspension of judgements, and hence, a fortiori, of prejudices, has logically the structure of a question. The essence of the question is the opening up, and keeping open, of possibilities” (Gadamer 1975: 266). This hermeneutical stance coincides with Max Weber’s intention to establish an interpretative sociology as a science of historical possibilities. To realise the objective opportunities in the historical data, we must distance ourselves from the given social reality and allow ourselves for that moment to become alienated from it. Or, as Helmuth Plessner put it, “the art of alienated perception fills therefore a necessary condition for all true understanding. [...] Without alienation there is no understanding” (Plessner 1983: 94).

Contrasting

The interpretation of the first sequence is finished when no new hypotheses on the meaning of the sequence can be formulated by rational means. At this point, this first probe will seem to be valid in relation to the sequence and perhaps to the research goal, but not necessarily in relation to the datum, i.e. the video-record as a whole. Consequently, the video-record as a whole must be examined in search of a second sequence that should contrast maximally to the first sequence and therefore offer a opportunity to falsify the first interpretation. For our case study, we contrasted a computer-animated sequence with a scene of face-to-face-interaction.

To widen the scope for new aspects or features of the case, we might look for other key scenes and repeat the procedure as long as all hypotheses on the meaning of the case can be reduced to a single one. Finally, when all elements of the material will have been reconstructed as moments that fit together and constitute a whole, the interpretation has to be considered as exhausted to the empirical data and valid as a theoretical model of the case in study.

The “GDR-Show” – A Case Study

Generally speaking, we do not treat video as an instrument to generate data, but instead understand it as a fully-fledged social construction of reality. This means that, even if we record data by video for scientific usage, we interpret them as ‘natural’ in the sense of Alfred Schutz’ concept of “first order construction” (Schutz 1953) – whatever the ‘real’ nature of the data might be. We will demonstrate this on the basis of the interpretation of two sequences from a video-record of a TV-show: a computer generated artefact on the one hand and a ‘real’ social interaction in front of the camera on the other hand.

From the Ostalgia-shows broadcasted by German television in the summer of 2003, we chose a video-record of the most successful Ostalgia-show on the private RTL network. The “GDR-Show” was seen by six and a half million Germans, or twenty percent of all TV-households in the country. It was the last programme in a series of similar shows in the retro-format presenting retrospectives of West German popular culture from the 1970’s and 1980’s. The popular culture of Eastern Germany was adapted to this blueprint. More or less prominent persons took the stage. Accompanied by media flashbacks, they reported their personal experiences in former times.



Image 1: Intro to the “GDR-Show”

The TV-format imposed by the presence of a presenter forced the guests of the show to distance themselves from their former lifestyles in an ironic way. The two following sequences from the “GDR-Show” demonstrate the ironical dissociation as a basic principle of the actors’ performances.

For a systematic analysis, we have chosen the introductory sequence to the “GDR-Show” as a first key scene. The introductory sequence appears at the beginning,

after every commercial break, at the end of the show, and functions like as a bracket that ties up all parts of the process of communication. During the show, the introductory sequence mutates into a shibboleth of the TV-serial and – as the following analysis will demonstrate – represents form and content of ‘Ostalgia’ in a nutshell. Thus, an interpretation of this short piece as a condensed model of ‘Ostalgia’s’ metaphorical content will give us insight into the structure of the phenomenon as a whole. This first hypothesis is contrasted by the sequential analysis of a second key scene from the show: a face-to-face interaction, which will be analysed with particular regard for the dialogues and for the body language of the presenter and a guest-star.

When the introductory sequence starts, we see a golden garland of corn emerging from darkness and moving from the right to the middle of the screen. A bundle of rays illuminates the scene, revealing a red background. The circle formed by the garland and its glow call a halo to the viewers mind. The manner in which the garland throws its shadow against the red background, and the rolling rays radiating from the centre behind the garland generate enchantment. Then moving letters appear on the screen which after a moment which arrange themselves to form the writing “Die DDR

Show” at the centre of the garland in a funny style. The use of moving letters as if they were pieces of a puzzle suggests that the previously established aura has been broken. All this lead us to infer a ritual. From this ‘in-between’ perspective, a state rooted in the interplay between numinosity and profanity is constituted. The modified GDR-emblem symbolises the virtual reality of the show. Right away, we cannot help but notice that the hammer and compasses, the core elements of the former East-German national emblem, are missing. The hammer, symbolising the working class and the compasses, representing the technocratic avant-garde, both elemental to the socialist state, have vanished. Only the garland of corn remains in a slimmed-down form as a reminiscence of the pastoral component of the GDR. Finally, the symbolic character of the emblem is depoliticised and replaced by a consumer good’s trademark.

At this point, the three letters break ranks, a spotlight falls on the middle of the letter ‘D’, and the camera dives into the inner blank space of this letter. At this moment of being ‘in-between’, the stage appears and the show starts. The symbol that had functioned as ‘a bridge to transcendent realities’ is transformed into a frame for a virtual sphere of action, thereby making the end of our analysed sequence.

During the sequence, we hear the “Prinzen”, a pop-band well-known both in the GDR and now in reunified Germany, singing “All this is Germany. All this is us” (“Das alles ist Deutschland. Das alles sind wir”). Each single phrases alone represents a statement. When the sentences are joined, they become normative imperatives, requiring identification of the viewer. After each sentence has been sung, it is immediately punctuated by the expressive phrase “oh-oh-oh”, which serves as a comment. The meaning of the “oh-oh-oh” is ambiguous. It might be an expression of admiration or an indication that these are delicate issues. This song is also heard at the beginning of



Image 2: World-star and presenter

the second key scene, in stilling it with a sense of ambiguity.

In the analysis of the first sequence, we identified a distinct pattern, constituted by three binary oppositions: (de)politicisation vs. commercialisation, auratisation vs. profanation, admiration vs. reservation. In the following, we attempt to show that these binary oppositions constitute the framework of the “GDR-Show” as an experimental setting in which the actors must

prove themselves as the avant-garde of a new collective identity construction. This was no easy undertaking, as, even in 2003, Eastern German traditions and mentalities were still regarded as a culturally archaic burden in the politically and economically reunified German Federal Republic. Both, the improvement of the actors and the construction of a collective identity will be realized if all the contradicting elements will fit together in one figure.

The former world champion of ice-skating and Hollyday-on-Ice-Princess Katharina Witt is introduced as an “Eastern-star”, a “Western-star”, and in the end as a “world-star” by the presenter. When the world-star descends the big steps of the stage wearing the uniform of a “Young Pioneer”, i.e. dressed as a member of the former communist youth organization, the audience is surprised. This discrepancy in the appearance of Witt is expressed by the ambiguity of the phrase “oh-oh-oh” in the theme song. The presenter refers to the incoherency in Witt’s appearance when he opens the talk with the remark “fesch”, which means, (am.) ‘fresh’ or ‘fine’, or (engl.) ‘bold’ or ‘keen’, but also ‘smart’, ‘elegant’ or even ‘sexy’.¹ By addressing her as ‘smart’ he expresses his admiration for Witt’s attractiveness, but also for her courage in wearing a costume that is taboo in reunified Germany. The ambiguity of the first comment is stressed by the second statement referring to the uniform (“Eine Jungpioniersuniform”), which manifests the presenter’s reservation respect to Witt’s outfit and its political correctness.²

This contradicting evaluation of Witt’s appearance by the presenter defines the limits of the space; the ‘in-between’ Witt must prove herself as a competent actor and ‘world-star’. She is confronted with the task of overcoming the contradiction of her appearance by performing a role in which the contradicting tokens ‘admiration’ and ‘reservation’ form one unit.

The audience’s reservation arise from the fact that the uniform is a communist symbol originally worn by children. Used as a prop in service of an adult women’s self-representation the uniform transforms itself into a fashionable accessory, thereby losing its political meaning. Witt confirms the presenter’s remark with the words: “Yes, I am here today as a Young Pioneer, so to speak” (“Ja. Ich komme sozusagen als Jungpionier heute”). In the German sentence the parenthesis “so to speak” (“sozusagen”) and the adverbial phrase “today” (“heute”) mark the distance to and the playful character of her role as Young Pioneer. She rejects an interpretation of the uniform as a symbol of uniformity, suggesting that she now lives in a ‘multi-optional world’. Nevertheless, her attitude of distance to the socialist past is counteracted by her posture, her gestures, and her facial expression. Sitting up straight, she proudly presents her scarf by pushing up the lappets. Finally, when she enunciates the word “Young Pio-

1 For the following analysis see the score in the appendix.

2 The choice of the presenter’s term “fesch” (engl. ‘smart’) as an expression for Witt’s attractiveness is motivated by her uniform. In German “fesch” is an idiomatic expression for the appeal of somebody in uniform.

neer" ("Jungpionier") with affection, she closes her eyes and, metaphorically spoken, outgrows beyond herself. In this way, her body-language indicates a regressive identification with her former role as a Young Pioneer and confirms the presenter's comment "fesch" ('smart'). By performing her personal identification with the role of a Young Pioneer, she solves two problems of self-representation: the contradiction between girl and woman and between being smart and appearing in a traditional communist outfit.

However, there is one element in Witt's performance, that still does not fit with the interpretation we have offered so far. Sitting up straight, Witt keeps her hands on her knees which are pressed together. This attitude of reservation is a reaction to the presenter's phrase "fesch" ('smart'), which is much less a comment on the outfit, but rather a compliment to the person wearing this outfit. The gesture is even a rejection of the sexual allusion contained within this compliment. Then, moving her hands to the scarf, she reacts to the presenter's second statement referring to the uniform and diverts his and the audience's attention away from her body to the accessory of the costume that defines it as a uniform. Proudly presenting the lappets of her scarf she brings out the attractiveness of the uniform and at the same time she gives the situation an erotic timbre.

Summarizing this scene we can say, that in a first step Katharina Witt reimbues a depoliticised uniform with infantile innocence, transforming it into a profane commercial commodity. In a second step, she bestows new charm upon this commodity. Marking it as the outfit of an idol of the international entertainment industry, Witt supplies the costume with new attractiveness. All three of the binary oppositions that we identified in the first sequence reoccur in the second sequence: (de)politicisation vs. commercialisation, auratisation vs. profanation, admiration vs. reservation. Here, reservation is still expressed with reservation.

Let's have another look at the scene. Although announcing reservation by putting her hands on her knees in reaction to the compliment and sexual allusion of the presenter, Witt confirms the erotic and sexual undertone in the interaction by her next gesture, that is by lifting the lappets, but – and this seems to be crucial – she defines the eroticism of the situation by *herself*. Generally speaking sexual attractiveness for an actor is a chance to become a media-star, but at the same time represents a high risk of losing one's face, one's authenticity, one's 'real' identity. The presenter's compliment puts Witt into a risky situation. She solves this problem opting for coquetry. She transforms the seduction of the presenter in his double role as a show-master and a man into an erotic interplay dominated by the woman. Coquetry is the reversal of the classical situation of seduction that follows the cultural definition of passivity as a female and activity as a male characteristic.

The seduction that a man tries to initiate as a representative of the hegemonial culture and in his role of a show-master is transformed into coquetry by the woman. By this redefinition of the situation, Witt's regressive identification becomes recognizable as a mask of her sovereignty and the attempt to prove herself as a self-confident woman *and*

as an Eastern German citizen as well. Following Georg Simmel (1996), coquetry is a state of 'in-between' the flirt and the beginning of the sexual intercourse. It is an erotic interplay between two equal antagonists of different gender. From this perspective the play of coquetry between Witt and Geissen is a parable for the newborn collective identity, which stands for an ironical state of 'in-between' Eastern and Western Germany.

Conclusion

Our example might demonstrate that *video hermeneutics* are not the application of a formal method on any object, but rather the representation of the social construction of reality within case studies. Collective identities are social constructions, generated by rituals as a specific kind of social action. In complex societies, it is not possible to fabricate generally binding constructions of collective identities through rituals requiring face-to-face interaction, the presence and participation of community. Today, only medial communication provides near ubiquity, reaching as many people as possible. In this way, the media achieve the condition for construction of collective identities and are bound only by the rules of media reception and not to political participation.

Our analysis has reconstructed the objective meaning of the media-product, not the aesthetic, ideological or economic intentions of the producers of the show. Neither these intentions of the producers nor the reconstructed objective meaning of the media-product has any social relevance. As symbolic orders in the sense of Ernst Cassirer (1965), the media are a reality of their own. Only the viewers, through their perception, give a pragmatic accent of reality to the fictitious medial constructions. Form and content of media-products gain historical relevance if the recipients attach value to them in reference to social action in everyday-life. As a consequence, *video hermeneutics* enable us to reconstruct the conditions under which media-products are designed. Furthermore, we can formulate not only well-founded hypotheses regarding the modes of perception and acceptance of these audio-visual constructions, but also with respect to their impact on social action in everyday-life.

References

- Berger, P.L. & T. Luckmann, 1966: *The Social Construction of Reality. A Treatise in the Sociology of Knowledge*, Garden City, New York: Doubleday
- Bergmann, J.R., H-G Soeffner, & T. Luckmann, 1993: Erscheinungsformen von Charisma – Zwei Päpste, in: W. Gebhardt, A. Zingerle, & M. N. Ebertz (eds.) *Charisma. Theorie-Religion-Politik*, Berlin: de Gruyter, 121-155
- Boehm, G. 1994: Die Wiederkehr der Bilder. In: G. Boehm (ed.) *Was ist ein Bild?*, München: Fink, 11-38
- Cassirer, E., 1965: *The Philosophy of Symbolic Forms. Volume 1: Language*, New Haven: Yale University Press

Fellmann, F., 1995: Innere Bilder im Licht des imagic turn. In: K. Sachs-Hombach (ed.) *Bilder im Geiste*, Amsterdam: Rodopi, 21-38

Gadamer, H.-G., 1975: *Truth and Method*. New York: Crossroad

Glaser, B.G. & A.L. Strauss, 1967: *The Discovery of Grounded Theory. Strategies for Qualitative Research*, Chicago: Aldine

Goffman, E., 1974: *Frame Analysis. An Essay on the Organization of Experience*, New York, Evanston, San Francisco, and London: Harper and Row

Goody, J., 1968: *Literacy in Traditional Societies*, Cambridge: University Press

Goody, J., 2000: *The Power of the Written Tradition*, Washington: Smithsonian Institution Press

Langer, S.K., 1967: *Philosophy in a New Key. A Study in the Symbolism of Reason, Rite, and Art*, Cambridge: Harvard University Press

McLuhan, M., 1964: *Understanding Media. The Extension of Man*, London: Routledge

McLuhan, M. & B. R. Powers, 1989: *The Global Village. Transformations in World Life and Media in the 21st Century*, Oxford: Oxford University Press

Mannheim, K., 1960 [1936]: *Ideology and Utopia: An Introductory sequenceduction to the Sociology of Knowledge*, London: Routledge

Mead, G.H., 1967: *Mind, Self, and Society from the Standpoint of a Social Behaviorist*, Chicago: University of Chicago Press

Mitchell, W.J.T., 1994: *Picture Theory. Essays on Verbal and Visual Representation*, Chicago: University of Chicago Press

Plessner, H., 1983: Mit anderen Augen, in: ders.: *Gesammelte Schriften VIII: Conditio humana*, Frankfurt a.M.: Suhrkamp, 88-104

Rorty R. (ed.), 1967: *The Linguistic Turn. Recent Essays in Philosophical Method*, Chicago: University of Chicago Press

Schutz, A., 1953: Common-Sense and Scientific Interpretation. In: *Philosophy and Phenomenological Research*, 14, 1-37

Schutz, A. & T. Luckmann, 1974: *The Structures of the Life-World, Vol. 1*, London: Heinemann

Schutz, A. & T. Luckmann, 1989: *The Structures of the Life-World, Vol. 2*, Evanston: Northwestern University Press

Simmel, G., 1996 [1918]: Die Koketterie, in ders.: *Hauptprobleme der Philosophie - Philosophische Kultur, Gesamtausgabe Band 14*, Frankfurt a.M.: Suhrkamp, 256-277

Soeffner, H.-G., 2004 [1989]: *Auslegung des Alltags - Der Alltag der Auslegung. Zur wissenssoziologischen Konzeption einer sozialwissenschaftlichen Hermeneutik*, Konstanz: UVK

Weber, M., 1978 [1921]: *Economy and Society. An Outline of Interpretative Sociology*, Edited by G. Roth and C. Wittich, Berkeley, Los Angeles, and London: University of California Press

Appendix: Score "DDR - Show"

| Our Time (sec.) | Visual Data | | | Audio Data | |
|-----------------|--|---|--|---|---|
| | Setting | Camera | Body | Head | Speech |
| 1 | Wirt and Geissen sitting on a modern sofa in front of a background, dominated by red and brown colours. Wirt wears a dark blue pleated skirt, which ends at her knees, a formfitting white blouse with epaulettes, and a knotted scarf. Her short braids are held together by rubber bands. | Knee shot slightly from the left. | Distance between the bodies 1 meter. Geissen's right arm outstretched on the sofa's backrest in direction to Wirt. His left arm resting on his thigh. Wirt sitting upright, both arms outstretched and hands on her closed knees. Wirt turning to the right and sitting up grasps and lifts the lappets of her scarf with both hands | Geissen stares at Wirt's blue scarf, smiling. Wirt turns her head to the right side. Wirt turns her head to the left side. Looking at Geissen. Geissen still staring at the scarf. | |
| 2 | Bright background: two bright spots right to Wirt's mouth. | Close-up view on Wirt, slightly from the right side; Wirt slightly right from the center of the picture; her mouth in the golden section. | Wirt plays with the lappets of the scarf. Wirt drops the lappets. She lifts her shoulders and stretches her back. Wirt drops her shoulders and unbends her back. | | Geissen: "Eine Jungpioniersuniform." Wirt: "Ja." Wirt: "Ich komme sozusagen als Jungpionier heute." Geissen: "Ja." |
| 3 | | | | | |
| 4 | | | | | |
| 5 | | | | | |
| 6 | | | | | |

Cornelius Schubert

Video Analysis of Practice and the Practice of Video Analysis Selecting field and focus in videography

When practicing videography, interrelations between the empirical field, the analytic focus and the technical instrument need to be taken into account. This article will address the issue of selecting data in various stages of research by using Grounded Theory's *theoretical sampling* as an exemplar of how to make selections that are both based on data and controlled by theory. The different aspects of selection will be illustrated by examples from a study about the work of anaesthetists in operating theatres.

Videography

Videography as a method is still young and no fixed rules exist as to how to go about filming and analysing real life situations. As a form of observation, videography has its ancestry in the ethnographic research tradition (cf. Mohn this volume), but video recordings are necessarily different in scope and focus from observations made with the naked eye. This article will be especially concerned with how scope and focus are related to theoretical assumptions, i.e. the apparent contradiction between classic ethnographic research, which insists on openness, and videography, which should be considered as a way of focusing. For this purpose, the method of *theoretical sampling* proposed by Glaser and Strauss (1967: 45) for selecting and comparing data in a process of controlled data collection will be illustrated by examples of videographic research conducted during an ethnographic study of work practices in a surgical operating room (OR).

The main question to be answered in considering the interrelation of video cameras and the sociological perspective is the question of how the researcher's knowledge of the field as well as prior theoretical assumptions are involved in the process of configuring the videographic focus. Glaser and Strauss state that theoretical sampling "is *controlled* by the emerging theory" (1967: 45) and based on general assumptions. Even though they reject the idea of using preconceived theoretical frameworks for making initial decisions of what to study, there is still room for *theoretical sensitivity* in the sense of the sociologist's "theoretical insight into his area of research" (ibid.: 46). Since the sociologist can neither shed all prior knowledge, nor should preconceived theoretical frameworks interfere with data collection, there remains only the solution of commencing research by explicating the theoretical research focus, while at the same time remaining open to the peculiarities of the field. To remain sensitive to both theory and field is a persistent facet of videographic research. In our case, the study was con-

cerned with the interrelation of human and non-human agency in the OR against a background of Science and Technology Studies, especially Actor-Network-Theory. We were especially interested in the distribution of activities between the OR-personnel and the OR-equipment, with respect to the routines and risks that such *distributed actions* (Rammert 2002) might pose.

It follows that because of the intention to highlight the interrelation of theory and data, the arguments in the following chapters will alternate between rather abstract theoretical considerations and very practical empirical explications.

Video-Analysis of Practice

The sociology of work in a wider sense has a long tradition in sociology in general (cf. Hughes 1971) and the work of doctors has drawn specific sociological attention from a number of scholars (eg. Freidson 1970, Strauss, et al. 1997[1985]). The ethnographic research of work sites and work practices has also been widely used in constructivist laboratory studies within Science and Technology Studies (Knorr-Cetina 1981, Latour & Woolgar 1979) and the ethnomethodologically orientated Workplace Studies (cf. Heath, Luff & Knoblauch 2004), the latter already making extensive use of video recordings. In the last two research areas, observations of practices are reconstructed in the light of the research question: laboratory studies set out to deconstruct the 'hard facts' of natural science and the workplace studies aim to reveal the social organisation of cooperation in high-technology work settings. In terms of Grounded Theory (GT), they remained theoretically sensitive with respect to their analytical background while producing naturalistic accounts of real world phenomena. In this way, observations can be conducted that are naïve, yet not so naïve in the way that they reveal the contingency and inconsistency of normal, every day routine action and interaction.

Filming in the operating room is different to 'simple' observation of daily practices. In addition to theoretical considerations, technical and ethical considerations shape the way in which practices can be recorded. First of all, access to the OR is controlled for technical and ethical reasons, making it largely inaccessible to the public. Secondly, because of the privileged access to the OR, there is little common knowledge about how work is done and in which ways people interact. Therefore, before videotaping highly skilled and specialised work practices, the observer has to accumulate a large amount of contextual knowledge in order to be able to understand what is going on. The ability to make sense of the observed procedures is an indispensable prerequisite to analysing work practices and prior participant observation lends itself well to gaining knowledge about the field.

As mentioned above, the study this paper draws on was conducted against the background of Actor-Network-Theory (ANT) proposed by Michel Callon (1986) and Bruno Latour (1988). Without going into the details of ANT, its contribution to study practices in the OR is the so-called *symmetry principle*, which states that human

and non-human activities should be considered as equal contributions in the analysis of an actor network. Therefore, both humans and non-humans are considered to be *actants* in an actor network and their relations are the relevant unit of analysis (Akrich & Latour 1992). By using this approach, two assumptions, one lay and one professional, that could obstruct the view of the interrelations between humans and non-humans in high-tech work situations, are replaced by the analytic perspective of ANT. The lay assumption is so-called *anthropomorphisation*, i.e. that users treat machines like human beings. In this case, doctors and nurses talk about the operating equipment in terms of its personality, especially in cases of dysfunction, attributing this to poor motivation or hostile attitudes of the equipment. The professional assumption is the sociological consensus that only human beings can really act, which translates to a (undoubtedly useful) *anthropocentrism* in sociological theory. Either way, the relation of humans and machines cannot be adequately described in terms of anthropomorphisation or in terms of anthropocentrism.

In order to observe, it is necessary to dispose of knowledge taken for granted. This is more troublesome than it sounds, since knowledge taken for granted has the virtue of blending itself into the routines of daily life. To observe humans and non-humans as equal in their contributions to collective work processes certainly distorts the assumptions taken for granted in either lay or professional perspectives, because artefacts are not usually perceived as an active element of cooperation. Thus ANT, in our case, is used not as a theory, but as a methodological instrument for epistemic purposes, putting the researcher in the position of a stranger. It is now possible to qualify the naïve, yet not so naïve, observations further. They are naïve in the sense that they need to be open to the phenomena of the field without a predetermined set of observation criteria, yet they are not so naïve, because a large amount of contextual knowledge needs to be generated first and the relevant analytical perspective needs to be constructed in order to frame the observations.

For the relation of participant observation and videography this means that the latter must follow the former. In addition, one must abandon the hope of being able to 'film it all'. Users will always find unobserved areas to communicate or shield relevant areas from scrutiny with their bodies (cf. Lomax & Casey 1998). The videographic perspective has a scope defined by the research interest, not by a desire for exhaustive documentation. Here, the importance of theoretical sampling for selecting scope and focus becomes clear. The knowledge generated by participant observation guides the scope of the camera and the researcher is able to define relevant areas of interest, both temporal and geographical, with respect to surgical operations.

This led in effect to the use of handheld video camera in this research project, which was mainly used in the beginning of the operation during anaesthetisation. In early stages of the research project we used two tripod mounted cameras in two opposite corners of the OR, but this setup turned out to be inadequate in capturing the rather small scale manipulations of machines and patients by the hands of doctors and nurses. Also, once the operation commences, surgeons gather around the incision, ef-

fectively blocking views from the peripheral areas of the room. The anaesthetist's workplace is also framed by the operating table and various machines block the scope of the cameras. A third aspect that led to us renouncing stationary cameras is the limited space available in operating rooms. Tripods are by no means allowed to stand in the way and sometimes it happened that while the mobile technical inventory was moved around in the OR, the tripods were accidentally displaced and the camera's angle shifted. Last but not least, the handheld camera can be quickly moved out of the OR, e.g. in critical situations when the feeling of being under surveillance might only put more stress on the staff. Thus the decision to switch from a stationary to a mobile camera was motivated by technical, ethical and practical considerations. But the use of a handheld camera is more selective than use of a stationary one. The mobile camera follows the researcher's gaze, highlighting certain situations while at the same time blinding out the rest.

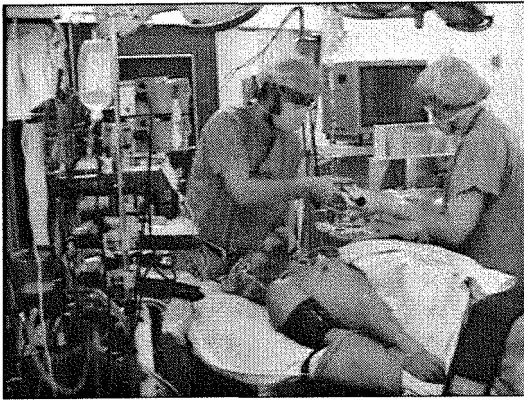


Fig. 1: Intubation by an experienced doctor

may be centred on the patient's body or may require a certain distance. If the operation becomes complicated, staff tend to move all marginal work away from the surgeons so as not to disturb them. A mobile camera can follow the doctors and nurses in action and in their movements through the OR and adjacent rooms.

Consequently, the video recordings are highly idiosyncratic. To overcome the limits of mere snapshots of interactions, one can follow Glaser and Strauss by *selecting comparison groups*. Since procedures in routine anaesthetisation are rather standardised, it is possible to define different stages and activities that can be compared with one another. In GT the groups are selected according to their *theoretical relevance*: group comparisons "are made by comparing diverse or similar evidence indicating the same conceptual categories and properties" (Glaser & Strauss 1967: 49). In the study of cooperative work sites, this often translates to comparing the interactions in mixed teams (humans and non-humans).

For example, we recorded a number of intubation sequences where either an experienced or a novice anaesthetist performed the procedure. During intubation (specifically

The video recordings produced by a handheld camera contain dramaturgic elements with respect to the composition of the recorded scenes. Zooming in and out are ways of scaling the focus from spoken human-human interaction down to the movements of a single hand manipulating instruments. The researcher can also follow the actors into the preparation rooms or to the telephone if sequences of interest are carried on outside the OR. Activities

transoral endotracheal intubation), a plastic tube is inserted through the patient's mouth into the trachea in order to administer artificial respiration. Depending on the shape of the tube and the patient's body, intubations can be more or less difficult to perform and if the anaesthetist misses the trachea, the food tube can be intubated by mistake. The ability to perform good intubations mainly comes from repeated practice, so novices need to be trained in the practice in order to become competent. For experienced doctors, the procedure is part of daily routine. In the scenes we recorded, the doctor is assisted by a nurse who hands over the instruments and takes them back. In figure 1 we see a nurse handing over the laryngoscope (an instrument for looking down the throat at the larynx) to the anaesthetist. This is an experienced team and the nurse has already turned her head away from instrument and anaesthetist to reach for the tube. An intubation usually takes between 30 to 90 seconds to perform and is supposed to happen quickly, since the patient does not breathe during this time. Such situations are well suited to videographic scrutiny, because they are locally confined and temporally finite, they happen without a lot of talk, and interaction is conducted swiftly. The details of cooperative work on such a small scale are hard to observe with the naked eye.

By the same token, the study of interactivity with machines can be appropriately recorded using video, making high-tech workplaces especially interesting for videographic studies (cf. Suchman 1987). The (mostly) nonverbal conduct when operating machines and the swift routine manipulations of artefacts can be better captured by video than by plain observation (cf. Schubert 2002). In our case, we concentrated on the use of the anaesthetic monitoring equipment and syringe pumps. Syringe pumps are microelectronic devices which assist in regulating

the administration of accurate amounts of drugs, such as sleeping or pain medication, over time. The interesting thing about syringe pumps is that they do not have the best reputation for being reliable. This is not to say that they cannot be depended upon at all, but their ability for self diagnosis can lead to false alarms and sometimes the digital readout can be misleading.¹ In

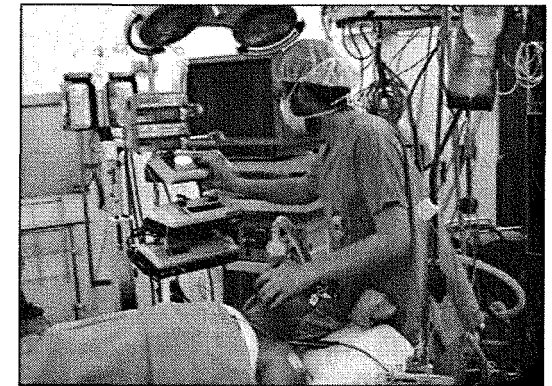


Fig. 2: Operating syringe pump

figure 2 we see an anaesthetist re-

1 A medical doctor concludes her review of a critical paper on syringe pumps thus: "With advances in technology, we tend to put too much faith in machines and their arrogant digital readouts. This paper jogs us back to reality." (<http://www.pedsanesthesia.org/newsletter/2002winter/litreview.shtml> - 22. June 2005)

programming a syringe pump after several internal alarms. In the daily use of these machines, anaesthetists develop specific routines for normal use and for troublesome situations. The latter can be seen as compensating activities (cf. Hughes 1951) that develop in the interaction with the machines and are connected with the local context. Since these activities are part of the routines taken for granted in everyday work, they are hardly reflected on by the users and the fine-grained documentation available in video recordings is well suited to making these routine practices available for analysis.

Selecting groups of sequences that are relevant to the research question is of course a process which is never absolutely finished, because the research question may shift in the light of new data. As a criterion for adding more sources of data or making an intermediary stop in the selection of other data, Glaser and Strauss propose the term *theoretical saturation*, which will be used for structuring and organising the practice of video analysis.

Practice of video analysis

Once the researcher has returned from the field with written observation protocols and video recordings it is time to analyse the data. One of the first steps in organising the data is to make *content logs* of the video recordings (Jordan & Henderson 1995, Ruhleder & Jordan 1997) in order to discover "easily identifiable behavioural units", or *ethnographic chunks* (Jordan & Henderson 1995: 57). Content logs are not transcripts of a complete tape, but rough descriptions of the filmed situations and may contain references to analytical concepts. The content logs change as the research progresses: they become more detailed when sequences are analysed, which can be seen as a similar process to that of *coding* in GT (Strauss 1987). The role of *theoretical saturation* in the process of identifying ethnographic chunks and creating content logs is that an analysis is always conducted with respect to the progress of the research. Glaser and Strauss define theoretical saturation as "the criterion for judging when to stop sampling the different groups pertinent to a category" (Glaser & Strauss 1967: 61), i.e. the researcher starts coding with theoretical sensitivity, continues to refine the categories by theoretical sampling and the process comes to an (sometimes tentative) end, when theoretical saturation is reached.² In the practice of video-analysis, this process resolves into the multiple steps and iterations of analysing videographic data, which are oriented towards the relevance of the material for the research question: a) selecting key sequences: looking at intubations with respect to novice/expert differences, b) repeated

² The different stages of coding (open, axial and selective) will not be elaborated here, because this would require a longer examination of the coding paradigm in relation with predetermined research questions. Suffice to say that it is a process of focusing and organising data guided by theoretical assumptions.

viewing: with slow or fast motion and c) systematically comparing different cases: from large and small hospitals in Germany and abroad for example.

At this point, it is important to note that after the ethnographer comes back from the field, the subsequent analysis is often conducted in groups. After the first step of creating a rough content log and identifying ethnographic chunks, a selection must be made to decide which sequences should be marked for further scrutiny. This can be helped by a collectively generated content log in two ways: firstly, research colleagues can make independent content logs of the same recording in order to broaden perspectives, secondly, they can successively revise an initial content log in an attempt to refine the existing observations, remarks and conclusions. Either way, for the analysis of practices in the OR this means that all the research colleagues must have some knowledge about the context of the situation, i.e. they must be familiar with the peculiarities of the field to some degree. By collectively creating a content log, the problem of individual biases can be countered and the analysis of video recordings can be subsequently conducted in groups.

For our study this was very important, because the research project was part of an interdisciplinary approach to safety in socio-technical systems. Together with colleagues from fields such as psychology, semiotics, ergonomics and computer science, we aimed at producing different accounts of the procedures in the OR according to each individual discipline and then unifying them in an abstract model of OR safety aspects. The video recordings served as a focal point for interdisciplinary research, since the establishment of a common data source (video recordings) made it easier to engage in interdisciplinary discussions of work practices.

Here, a technical benefit of video recordings becomes evident: the capacity to store otherwise volatile interactions, gestures, etc. and the suitability for posterior analysis (cf. Mead 1975). Once a content log is established and the relevant scenes are agreed upon, the detailed scrutiny of selected sequences can commence. Each sequence will be analysed,

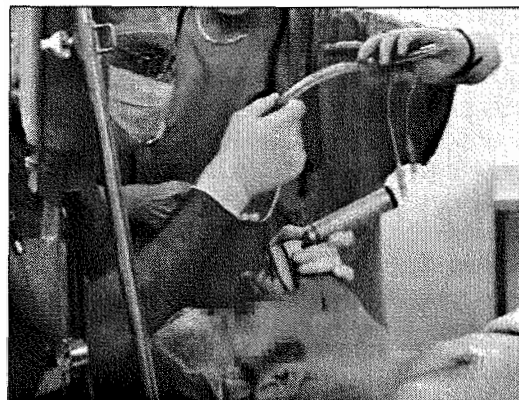


Fig. 3: Intubation by a novice doctor

i.e. repeatedly watched and discussed, until a sufficient level of theoretical saturation is reached. When analysing these sequences, a second technical benefit of video recordings comes into play: the possibility of distorting and altering the temporal composition of perception. By means of playback, a single scene can be endlessly repeated and slow motion and fast motion may reveal phenomena that are hard to perceive in real time.

For instance, the subtle coordination of team interaction (cf. Hindmarsh & Pilnick 2002) during the intubation could only be made visible by slowing the playback speed, sometimes even stopping the film and analysing the interaction frame by frame. In our example, we compared the interaction of experienced nurses with either trained or novice anaesthetists. In the first case, the nurse waited for the signals for coordination from the doctor, which were given in the form of swift gestures. Thus, each sequence in the interaction was initiated by the anaesthetist, which points to two interesting aspects of OR cooperation. First, the nurse recognises the doctor's authority concerning the procedure and takes on the role of an assistant and the traditional hierarchy of doctors and nurses is maintained, secondly, the nurse displays a high degree of competence, since she is able to react appropriately to even the slightest cues, like a quick nod or swift hand movement. The nurse assisting the novice doctor does basically the same, but she does so before the doctor signals her. By presenting the relevant instruments and nodding towards the patient, she is giving cues to the novice as what to do next. She arranges the instruments and tools in a way that makes it easier for the inexperienced doctor to successfully finish the intubation. By manipulating the material environment, she is able to instruct the novice by assisting him without having to use words. In figure 3 she is handing over the tube without letting it go as the novice doctor reaches for it, thus making him aware of the special position the tube needs to be in before insertion. Again, the traditional hierarchy is maintained, but only on a superficial level, with the nurse skilfully coordinating the interaction in a way that is barely visible to the naked eye.

The detailed analysis of such short scenes by means of video is obviously well suited for the task and it is the patterns of interaction and the adjustment of actions in problematic situations that especially render themselves visible through videography. This understanding of videography links it not only to ethnography but also to ethnomethodology (Garfinkel 1967, cf. also *video-interaction-analysis* Knoblauch 2004). The scope of video adds to that of participant observation by focusing on relevant scenes which are selected according to theoretical sampling and analysed until theoretical saturation is reached.

But video has one more advantage that should be noted: video recordings can be returned to the field in order to generate feedback, making video, in essence, a *medium for reflection* (Suchman & Trigg 1991). Feedback strategies are well known to ethnographers and visual techniques, e.g. *photo elicitation* (Harper 1984), are part of the repertoire of visual sociology. In our study we conducted *video-based interviews* with anaesthetists, using short scenes as a stimulus for comments and discussions. For the analysis of practices in the OR, this is relevant especially in validating the sociological interpretations of medical work and generating new perspectives on the data. The anaesthetists can assess whether a situation is typical for their everyday work and give background information about certain procedures or rate the difficulty of the task. Although the video sequence functions as a stimulus, the following discussion between sociologist and anaesthetist makes it different to Merton's *focused interview* (Merton & Kendall 1946). The aim is not to study the reaction to a stimulus, but to use the stimulus as a

starting point for a problem-centred talk, in which sociologist and medical doctor engage in the reconstruction of the scene from social, technical and medical perspectives.

In addition to validating sociological accounts of medical procedures and introducing a medical perspective into the analysis, we were interested in the way that these video sequences initiated oral stories of similar situations. The importance of *war stories* has been noted before (Schulman 1996, Rochlin 1999, Orr 1986, Barley & Orr 1997) and during the participant observation it was often the case, that when an experienced doctor instructed a novice, anecdotes of critical situations were used to highlight the central problems and to supply the novice with a memorable story concerning problematic situations. When the doctors were asked to describe the situation seen on video and to assess its potential for complications they very often interwove short episodes of critical incidents that they themselves or other colleagues had witnessed. These might range from patients waking during the operation or the breakdown of electronic instruments when transporting a patient, say, in an elevator. These anecdotes serve to make the novice more alert to practical hazards or undocumented interactions, e.g. between technologies and drugs. This rich corpus of anecdotes, episodes, war stories and practical accounts is what generally makes up the narrative structure of medical knowledge (Atkinson 1995, Hunter 1991). That these narratives are also produced in interaction with sociologists in discussing video sequences, indicates their widespread use in medical settings.

Thus video can be used not only for refined analysis, it can – in the sense of a medium for reflection – be a source of new kinds of data, i.e. the accounts given in the interviews. Videography generates multiple forms of pictures, sound and text, which, according to the principles of GT can be considered as *slices of data* (Glaser & Strauss 1967: 65). Within theoretical sampling, slices of data are “different views or vantage points from which to understand a category and to develop its properties” that help to exceed the limited input of only one method. Although most data may be gathered using one primary method, the process of analysis must contain elements of openness. Glaser and Strauss point out that slices of data should not be considered “as accurate evidence for verifications” (ibid.: 66), since “no data is accurate” (ibid.: 67), but should be seen instead as rich sources of information in the process of establishing categories. This does not contrast with the verifying of sociological accounts of medical procedures by doctors. Rather it is a sort of tentative *triangulation* (Denzin 1989 general see pp. 25, with respect to film see p. 225) which, unlike geometrical triangulation, does not try to identify an ‘objective truth’ of some kind, but attempts to view the phenomenon of interest from different angles, be they different theories, methods, researchers or slices of data.³

The previous remarks indicate that the practice of video analysis within videographic research is not confined to a single method of analysis. By the same token, a

3 An instructive example of how real world phenomena can be categorised with respect to analytic categories from several different perspectives is the case of *machine work* on hospital wards by Strauss et al. (1997[1985]: 40).

plurality of methods is not advocated for its own sake, but always with respect to theoretical sampling and saturation. Videography can therefore never be a monolithic method, applicable irrespective of the context of the study, but instead is a tentative approach that displays sensitivity both to the theoretical and to the empirical world. The practice of video analysis is thus a process of constant interrelations of theory and field.

Conclusion

The videographic approach put forward in the previous pages aimed at providing a theoretically orientated strategy for selecting and analysing audiovisual data in the continuous process of research. In this vein, there is no single optimal way to implement videography, it needs to be adapted to the respective study. The quality of videographic research does not lie in following strict methodological rules, but in the ability to arrange different methods, slices of data and theoretical assumptions in order to gain a deeper insight into the interrelations of real world phenomena. This is not, however, an invitation to arbitrariness. The researcher has to account for every move and measure in the research process on the basis of theoretical and empirical rationales. Videographic methods may differ for the analysis of school room interactions for example, in contrast to the methods used to observe operating rooms or auctions, laboratories, control centres, etc. Bearing videography's ethnographic tradition in mind, the general requirement is that it remain sensitive to the peculiarities of the field, whereas Grounded Theory additionally requires us to stay theoretically sensitive as well.

The central mode of operation in videography therefore consists of making distinctions and comparisons, iteratively introducing order into the 'messy' phenomena of the empirical world. The video camera and video playback equipment serve as instruments that render specific properties of observable phenomena visible. In contrast to the naked eye, the more confined scope of the video camera makes us aware of the selections we have to make during observation, but video recordings in return offer a very rich corpus of data for detailed analysis. For this reason, video recording and analysis in videographic research should be considered focusing devices which are embedded within a larger context of multiple methods, ranging from participant observations to interviews and producing very detailed accounts of selected phenomena in the field. Using video equipment as a sociological instrument, one has to keep in mind that it does not produce or reproduce 'reality' but that it consists of an array of artefacts which aid in the sociological reconstruction of practices by distorting our perceptual habits and exempting us from some restrictions of space and time.

The possibilities that videographic research offers have not been exploited to their full extent, and critical issues have not been completely solved. In terms of Grounded Theory, this means proposing videography as a flexible research strategy, rather than as a monolithic method, keeping it open for further refinement and modification. The

combination of videography with GT's theoretical sampling provides a lead for structuring qualitative research of work practices and the like.

References

- Akrich, M. & B. Latour, 1992: A Summery of a Covenient Vocabulary for the Semiotics of Human and Nonhuman Assemblies. In: W. E. Bijker (ed.) *Shaping Technology/Building Society. Studies in sociotechnical change*. Cambridge, MA: MIT,
- Atkinson, P., 1995: *Medical Talk and Medical Work. The Liturgy of the Clinic*. London: Sage
- Barley, S. R. & J. E. Orr (ed.) 1997: *Between craft and science. Technical work in U.S. settings*. Ithaca: IRL Press
- Callon, M., 1986: Some Elements of a Sociology of Translation: Domestication of the Scallops and the Fishermen of Saint Brieuc Bay. In: J. Law (ed.) *Power, Action and Belief: a new Sociology of Knowledge?* London: Routledge and Kegan Paul, 196-233
- Denzin, N. K., 1989: *The Research Act in Sociology*. Chicago: Aldine
- Freidson, E., 1970: *Profession of Medicine. A Study of the Sociology of Applied Knowledge*. New York: Harper and Row
- Garfinkel, H., 1967: *Studies in Ethnomethodology*. Englewood Cliffs: Prentice-Hall
- Glaser, B. G. & A. L. Strauss, 1967: *The Discovery of Grounded Theory*. Chicago: Aldine
- Harper, D. 1984: Meaning and Work: A Study in Photo Elicitation. *International Journal of Visual Sociology*, 2, 1: 20-43
- Heath, C., P. Luff & H. Knoblauch, 2004: Tools, Technologies and Organizational Interaction. The Emergence of Workplace Studies. In: D. Grant, C. Hardy, C. Oswick & L. L. Putnam (ed.) *The SAGE Handbook of Organizational Discourse*. London: Sage, 337-358
- Hindmarsh, J. & A. Pilnick 2002: The Tacit Order of Teamwork. Collaboration and Embodied Conduct in Anesthesia, *The Sociological Quarterly*, 43, 2: 139-164
- Hughes, E. C. 1951: Mistakes at Work. *Canadian Journal of Economics and Political Science*, 17, 320-27
- Hughes, E. C., 1971: *The Sociological Eye: Selected Papers*. Chicago: Aldine
- Hunter, K. M., 1991: *Doctors' Stories: The Narrative Structure of Medical Knowledge*. Princeton: Princeton University Press
- Jordan, B. & A. Henderson 1995: Interaction Analysis: Foundations and Practice. *The Journal of the Learning Sciences*, 4, 1: 39-103
- Knoblauch, H. 2004: Die Video-Interaktions-Analyse. *Sozialer Sinn*, 1: 123-138
- Knorr-Cetina, K., 1981: *The Manufacture of Knowledge: An Essay on the Constructivist and Contextual Nature of Science*. Oxford: Pergamon Press
- Latour, B., 1988: *The Pasteurization of France*. Cambridge, Mass.: Harvard University Press
- Latour, B. & S. Woolgar, 1979: *Laboratory Life: The Social Construction of Scientific Facts*. London: Sage
- Lomax, H. & N. Casey, 1998: *Recording Social Life: Reflexivity and Video Methodology*. Sociological Research Online
- Mead, M., 1975: Visual Anthropology in a Discipline of Words. In: P. Hockings (ed.) *Principles of Visual Anthropology*. Paris: Mouton, 3-12
- Merton, R. K. & P. L. Kendall 1946: The focused Interview. *American Journal of Sociology*, 51, 6: 541-557

- Orr, J. E., 1986: *Narratives at work: story telling as cooperative diagnostic activity*. Austin: ACM Press
- Rammert, W., 2002: Technik als verteilte Aktion. Wie technisches Wirken als Agentur in hybriden Aktionszusammenhängen gedeutet werden kann. In: K. Kornwachs (ed.) *Technik - System - Verantwortung*. Münster: LIT Verlag, 219-231
- Rochlin, G. I., 1999: The social construction of safety. In: J. Misumi, B. Wilpert & R. Miller (ed.) *Nuclear Safety: A Human Factors Perspective*. London: Taylor & Francis, 5-23
- Ruhleder, K. & B. Jordan, 1997: Capturing Complex, Distributed Activities: Video Based Interaction Analysis as a component of Workplace Ethnography. In: A. S. Lee, J. Liebenau & J. I. De Gross (ed.) *Information Systems and Qualitative Research*. London, UK: Chapman and Hall. London: Chapman and Hall, http://alexia.lis.uiuc.edu/~ruhleder/publications/97_IFIPWG82.html
- Schubert, C., 2002: *Making interaction and interactivity visible. On the practical and analytical uses of audiovisual recordings in high-tech and high-risk work situations*. Berlin: Technische Universität, Institut für Soziologie
- Schulman, P. R. 1996: Heroes, organizations, and high reliability. *Journal of Contingencies and Crisis Management*, 4, 2: 72-82
- Strauss, A. L., 1987: *Qualitative Analysis For Social Scientists*. Cambridge: Cambridge University Press
- Strauss, A. L., S. Fagerhaugh, B. Suczek & C. Wiener, 1997[1985]: *Social Organization of Medical Work*. New Brunswick: Transaction
- Suchman, L. A., 1987: *Plans and Situated Actions. The Problem of human-machine communication*. London: Cambridge University Press
- Suchman, L. A. & R. H. Trigg, 1991: Understanding Practice: Video as a Medium for Reflection and Design. In: J. Greenbaum & M. Kyng (ed.) *Design at Work: Cooperative Design of Computer Systems*. Hillsdale, NJ: Lawrence Erlbaum, 65-89

Anssi Peräkylä & Johanna Ruusuvuori

Facial Expression in an Assessment

Facial expression in psychology

More than 130 years ago, Charles Darwin (1872) pointed out that facial expression serves important adaptive function in regulating the interaction between humans (and many animals alike). He demonstrated how the expression of major emotions in face and body “could be analysed in terms of adaptive behaviour patterns, of which they were considered to be the rudiments” (Scherer 1996, 286). For example, the ways in which humans express negative feelings having to do with disgust and contempt through facial movements around the mouth and nose, and through turning away, may originate in specific situations where our ancestors have encountered offensive odours which they have tried to expel or exclude (Darwin 1872: 253-277).

Darwin’s work got forgotten for almost a century. In early and mid 20th Century, the study of facial expression was not central in the agenda of human sciences. However, in 1960s and 70s, Darwin’s work was found again, most notably, perhaps, by Paul Ekman who was involved in a cross-cultural study on facial expression of emotion (e.g. Ekman et al. 1969). Using photos and films of faces with different expressions, Ekman and his colleagues tried to pin down the connections between emotional states and details of the muscular movement in the face, as well as the ways in which people recognize such movements as expressions of particular emotions (for an accessible overview, see Ekman 2003; see also Izard 1971). Although Ekman discusses the uses of facial expression in social interaction (1979), the main focus of his work lies elsewhere, that is, in the ways in which internal emotional states are expressed and recognized in and through the face.

A rather different take on facial expression can be found in the work of Chovil (1991, 1997) and Bavelas (Bavelas & Chovil 1997; 2000). Rather than focussing on the functions of face as an output of internal emotional processes, they examine facial displays¹ as “visible acts of meaning”, by considering the ways in which the facial displays “are part of the integrated message with words” (Bavelas & Chovil 2000: 166; cf. Fridlund 1996). Using video recorded data from two-party conversations in a psychology laboratory setting, Chovil (1991) found two major types of facial displays. *Syntactic* displays involve facial expressions (most often, raising or lowering eyebrows) that serve for example in emphasizing or underlining what is said, or mark a question, or the beginning or the continuation (after a side track) of a story. In *semantic* displays, the facial expression for

1 Bavelas and Chovil use the term facial *display* rather than facial *expression*, thereby emphasizing the active uses of face in interaction (cf. Parkinson et al. 2005, 177). In this chapter, we will use the two terms interchangeably.

- Heritage, J. and G. Raymond, 2005: The terms of agreement: Indexing epistemic authority and subordination in talk-in-interaction. *Social Psychology Quarterly* 68:15–38
- Izard, C., 1971: *The face of emotion*. New York: Appleton-Century-Crofts
- Kendon, A., 1990. Conducting interaction. Patterns of behavior in focused encounters. Cambridge: Cambridge University Press
- Motley, M.T. and C. T. Camden, 1988: Facial expression of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication* 52: 1–22
- Parinson, B, Fischer, A.H. and Manstead, A.S.R., 2005. *Emotion in social relations. Cultural, group, and interpersonal processes*. New York: Psychology Press
- Peräkylä, A., 2004: Conversation analysis. In C. Seale, J. Gubrium, G. Gobo and D. Silverman (ed.) *Qualitative research practice*. London: Sage, 165–179
- Pomerantz A., 1984: Agreeing and disagreeing with assessments: some features found in preferred/dispreferred turn shapes. In J.M. Atkinson & J. Heritage (ed.) *Structures of Social Action*. Cambridge: Cambridge University Press, 57–101
- Ruusuvuori, J., 2001. Looking means listening: Coordinating displays of engagement in doctor-patient interaction. *Social Science & Medicine* 52: 1093–1108
- Schegloff, E.A., 1979: The relevance of repair to syntax-for-conversation. *Syntax and semantics* 12:261–286
- Scherer, K.R., 1996: Emotion. In: M. Hewstone, W. Stroebe & G.M. Stephenson (ed.) *Introduction to social psychology*. Oxford: Blackwell
- Sorjonen, M.-L., 2001: *Responding in conversation. A study of response particles in Finnish*. Amsterdam: John Benjamins

Monika Wagner-Willi

On the Multidimensional Analysis of Video-Data Documentary Interpretation of Interaction in Schools¹

Since the 1980s video-based research has enjoyed an increasing popularity in qualitative social research. This is not only true of sociological research but also for studies in other social science disciplines. Educational science, for example, in particular research into schooling and childhood, uses a video-based approach (see Huhn et al. 2000, Brandt, Krummheuer & Naujok 2001). Despite the broad scope of this method, methodological reflection into video-based research – not only with regard to the specific quality of audiovisual recordings but also with regard to the different ways of interpreting video data – is still in its infancy (see Knoblauch 2000, Wagner-Willi 2001).

In my contribution I will first address some methodological aspects of video analysis and then illustrate the procedural approach of documentary video interpretation using an empirical example taken from my study of schoolchildren's rituals. In conclusion, I will explain the practical uses of this method's multidimensional microanalysis.

Sequential Structure and Simultaneity in Video Data

The current lack of reflection upon video-based research stands in marked contrast to the development of textual interpretative procedures. Since the 1970s they have been consistently elaborated in the course of methodical and methodological reflection – e. g. the narrative interview and narrative analysis created by Fritz Schütze (see 1983, 1987). Up until the present, textual interpretative methods have dominated qualitative research. This predominance can be found in many video-based studies, e. g. in school research, where transcribed verbal communication is *central* to the analysis (see, e. g., Stadler, Benke & Duit 2001).

As Ralf Bohnsack (2003b: 241) makes clear in his critical diagnosis of qualitative methods of picture interpretation, the dominance of textual interpretative procedures relies upon the premise that social reality has to be presented in textual formats in order to gain scientific relevance. The premise is stretched to the extent that the “raw” data must also be in the form of a text, and, indeed, preferably a text created by the subjects of the research themselves. This methodological understanding of scientific knowledge was then drastically taken to be social reality by objective hermeneutics – in the sense that the world is to be understood *as* a text (see Garz & Kraimer 1994).

¹ Many thanks to Simon Garner for his great help in translating the text.

Against this background Martina Leber and Ulrich Oevermann (1994: 386) argue, that “the actual sequential structure of all social, psychological and cultural phenomena, objects and events” requires a sequential analysis.

This perspective of social reality as being primarily sequentially structured neglects, however, a further fundamental mode of structuring socio-cultural phenomena, namely, that of simultaneity. This mode is already demonstrable on a textual level, e. g. in transcripts of audio recordings of everyday conversations (see Sacks, Schegloff & Jefferson 1978). In these, notations are used: e. g., to show where verbal utterances overlap each other or to note an expressive activity, such as laughing, that accompanies the words. It is also here, where such audible activities are merely transcribed, that a mode of structuring is discernable, which I would like to term the *interweaving of sequentiality and simultaneity*. Methodologically, that means that a dominant premise of qualitative research – the assumption of a primarily sequential structure of social actions and interactions – must be respectively revised or modified.

The structure of simultaneity varies in its form and intensity depending on the specific situation. In this regard, different levels can be distinguished: (a) the social situation, (b) the interaction, and (c) the actions of individuals.² Levels (a) and (b) can merge into one, as, for example, in a therapeutic conversation; however they don't have to. During a school playground break we find the coexistence of a great number of interactions. On the level of interaction (b) we come across bodily, gestural-facial and verbal correlation. The level of individual acting (c) is characterized by bodily-sensual-spatial and verbal coordination. Thus we find simultaneity with regard to coordinated, correlated and coexistent activities. It is based on corporality, materiality, image quality and scenic arrangement – in other words on the *performativity* of social reality (see Wulf, Göhlich & Zirfas 2001).

While the basic model of sequentiality is a successively ordered text, the basic model of simultaneity is a picture. For, as the art historian Max Imdahl (1996: 23) makes clear, it is the “simultaneous structure” that is fundamental to the latter; a “sense-creating contemporaneity”.

So in social situations we find a performative process that is characterised by the interweaving of simultaneity and sequentiality. Video is particularly suited to recording this performative process. True, this procedure transforms the space into a two dimensional picture, however it retains the iconicity, i. e. the materiality or quality of image, and the simultaneous structure of social situations. A method for analysing video recordings should take account of this particular quality, i.e. it should not only be directed at analysing the sequential but also the simultaneous structure.

² Levels (a) and (b) correspond to Goffman's terms “social situation” and “encounter” (see Goffman 1981).

The Study

In my empirical study of children's rituals in primary school, I was searching for a way to analyse both sequentiality and simultaneity. The procedure was based on the method of documentary interpretation, conceptualized by the sociologist of knowledge Karl Mannheim and developed by Bohnsack. This investigation is embedded in a research project, lead by Christoph Wulf (see Wulf et al. 2001) focussing on rituals in different educational fields and forming part of a bigger interdisciplinary research project (Sonderforschungsbereich), entitled “cultures of performativity” at the Free University Berlin.

The aim of my study was the reconstruction of ritual practises and interaction of 4th to 5th grade schoolchildren, taking place in the classroom during the everyday transition from break to lesson (see Wagner-Willi 2005). These rituals were analysed with regard to their process and their social context of meaning. Thereby, the bodily, material, scenic and, in the sense of Wulf (1998), mimetic shaping of rituals, i.e., their micro-processual performativity was of particular interest.

Metatheoretically, the investigation was referring to Mannheim's (1980) concept of *conjunctive experiential space*. As a fundamental sort of sociality, conjunctive experiential spaces arise where people in the course of similar social experience develop shared forms of acting and shared forms of knowledge. This sort of knowledge is atheoretical and intuitive. Those who belong to the same conjunctive experiential space or (so to speak) cultural milieu *understand* each other *immediately*; they act and interact in *habitual concordance*.

Different from this form of sociality is the *communicative relation*. There, the participants have to *interpret* each other because they don't have a shared basis of experience. We find communicative relations in interaction contexts where people with different conjunctive experience are involved, in particular in institutional settings. While, for example, peer groups at school can be classified as conjunctive experiential spaces, lessons are constituted by a communicative relation context with different institutionalized social roles.

My study was directed towards the *How* of the ritual-interactive forms children show in the transition from break to lesson, a process in which the detachment from the peer group and the assumption of the role-like mode of acting of a schoolchild is expected by the institution. Thereby, the intermediate stage inside the classroom before the lesson starts was of special interest. In the sense of the anthropologist Victor Turner (1969), this intermediate stage can be understood as a *threshold* or *liminal phase*. It seemed to clearly highlight the area of tension between the peer group culture and the institutional order of the school (see Göhlich & Wagner-Willi 2001). Because the focus of the study was to investigate the performativity of rituals, it was appropriate to use a video-based approach.

For the audiovisual recording of the threshold phase between break and lesson in three classes, we used a fixed position for the camera. The camera was positioned so

that, as well as the threshold space of the door, a wide field of view of the room was also included. The videotaping was preceded by an explorative phase with participant observation.

Actionism in Classroom

The following analysis centers on scenes from a videotape that was recorded in the classroom of 4th grade schoolchildren, showing a phase in the transition from the first long break to lessons. In the analysis we will focus on two participants: Robert and Roswitha. (Video footage from 18.3.99, 10:26 – 10:32).

Territorial encroachments between Roswitha and Robert

Robert (not wearing his coat) enters the classroom after Andre and Michail. He walks rhythmically, briefly circling his lower arms around each other, to his seat near the door, closely followed by Roswitha (who is wearing her jacket open). As she passes, Roswitha places her hand on his back, pushes his upper body over the desk and presses him onto the desktop. She bends over him and slaps him several times, quickly and hard, on the back. Robert, who in being bent over grasps the thermos flask that is standing on his desk, cries out: "Ah, ah, ah, ah," as Roswitha slaps him. Roswitha straightens up again, thrusts out her chest as she walks, tosses her long hair from side to side and laughs. She turns towards Robert and takes off her jacket. Robert looks at her and holds the bent over pose for a while. Then he straightens up. As she takes off her jacket at her place at the neighbouring desk, Roswitha turns to face Robert again. Robert moves the items on his desk to the side. Then he jumps onto the chair in front of him and draws himself up to his full height. More and more children stream into the classroom. Robert casts his gaze around the classroom as he moves rhythmically again, rolling his hands in small circles around each other, and calls out in a sing song voice: "Ferero Küßchen". Meanwhile Roswitha goes back towards the door again with her jacket in her hand. She approaches Robert and looks at him. Just as she passes his desk, Robert jumps up onto the desktop, then down onto Sirin's (Roswitha's neighbour's) chair, over onto Roswitha's chair, back down onto the floor and runs to the window (on the other side of the room and out of camera view). Roswitha watches Robert's scramble over the desk and chairs and shouts: „Hey". Then she goes to the coat rack and hangs up her jacket. Sirin, who is approaching from the door and taking off her jacket, looks briefly at Robert, bends immediately over her chair and wipes the seat with her hand. Then she finishes taking off her coat and goes to the coat rack.

In the meantime, Roswitha has hung up her coat. (...) While the teacher, followed by Roswitha, steps into the aisle, Robert returns from the window area. When he arrives at Roswitha's seat he skips, supporting himself with a hand on the back of her chair. The teacher walks past Robert. Robert who has moved away from Roswitha's chair, is jostled by Roswitha as she returns. He bangs against Sirin's chair as he tries to support himself on the

desktop with his left hand. Roswitha grins broadly. Robert gets up and moves away while Roswitha takes her seat.

(After circa a minute:) Robert turns away from Binol with whom he had been speaking. With one foot on the chair, he leans his upper body over the desk, looks up and calls out loud: "Ferero Küßje, aah".

When Robert, with his dance-like motion, heads for his seat near the door at the beginning of the scene, he is caught up in a pointedly physical interaction by Roswitha. From behind, she bends his upper body over the desk top and slaps him repeatedly on the back. In a mutual interplay with these slaps Robert makes paraverbal noises: "ah ah ah ah", noises that sound joyful rather than upset and that hint at the sexualised meaning of

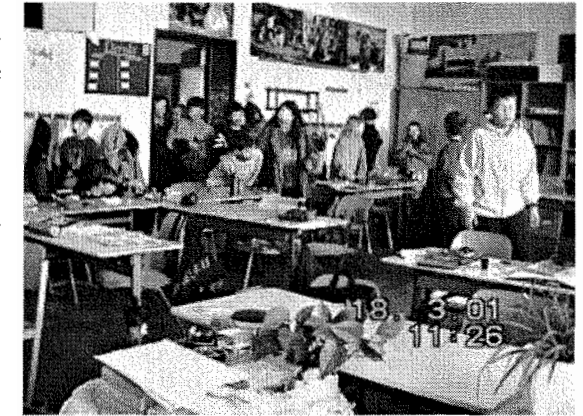


Image 1: Roswitha slaps Robert on the back

this physical encroachment, an encroachment initiated "from behind" in an "on-top/underneath" position. After Roswitha has finished with Robert, she makes her way to her desk nearby, thrusting out her chest coquettishly, swinging her hair and laughing. Her display gives the impression of being both triumphant and appealing. By provoking intense physical contact in a dominant position, Roswitha has deviated from the traditional gender roles which were most commonly observed in the video recordings. According to these, the offensive initiation of interactions between boys and girls is performed by the male party.

In the continued course of the interaction, Robert reverses the unconventional performance of roles mimetically. At first, with his jump onto his chair, he brings himself into a spatially higher and blatantly territorial position. As at the beginning, he makes a rhythmic dance-like movement again, but this time accompanied by a singing: "Ferero Küßchen" while Roswitha comes towards him. As she passes his desk, he seems to be running away from a further encounter with her. However, his flight includes a provocative territorial encroachment on the two girls' chairs. They show differing reactions: while Sirin immediately turns her attention to cleaning her chair, Roswitha stares at Robert in surprise and expresses displeasure ("Hey"). With his scramble over the desk and chairs Robert intensifies the dramaturgy of the whole scene: the movements become faster and the territorial encroachments more extensive and con-



Image 2: Robert scrambles over the desk and chairs

desk. She clearly exhibits her enjoyment in doing so with her grin. Roswitha's jostling is a renewed reversal of Robert's territorial encroachment on the chairs. At the same time she makes another physical encroachment. The gestalt of the scene is thus drawn back together both spatially and physically.

Not only a completion of the gestalt but also a sort of epilogue to this reciprocal and interactive entanglement is then offered to the observer: about a minute after it has ended Robert bends his upper body forward over his desk, looks up and says again loudly "Ferero Küssje... aah." Thus he repeats the physical-spatial positioning which Roswitha put him in at the beginning of the scene, and refers – by repeating the brand name of a praline chocolate – to the elements of enjoyment, pleasure, and eroticism that he once again articulates paraverbally as he did during Roswitha's initial encroachment.

The interactive involvement between Roswitha and Robert clearly displays traits of *actionism* in Bohnsack's (2004) sense, that is: a joint spontaneous action; an experimental quest for habitual concordance. These actionist practices present a particular form of dealing with differences in the liminal phase of the transition from break to lesson, and can be described as *conjunctive ritualisation* (Bohnsack 2004). This ritualisation fulfils two functions: 1. the processing of gender differences and 2. the processing of the difference between the rule structure of the peer group and that of lessons.

First this actionism is directed towards the experimental initiation of new relationship structures between the genders. This takes place in a spontaneous game of entanglement with the body and with the space and its inner territorial boundaries. The physical contact and the recurrence and reversal of asymmetrical positioning are just as much elements of this actionist quest as the sexualised gestures, dance movements and utterances.

The liminal phase of the transition from break to lesson is particularly suited to such actionist practices precisely because the conjunctive relationship structure created by

spicuous. They also include other children, who, as a background foil, highlight the mimetic relationship between Robert and Roswitha.

When Robert returns shortly afterwards, the gestalt of the mutual game with Roswitha is completed. His skipping is accompanied by a renewed territorial encroachment upon Roswitha's chair. Roswitha, who has just got back from the coat rack, then shoves him so that he bangs himself against the

the peer group in the break, with its predominantly gender-homogenous alignment (see Thorne 1993, Wagner-Willi 2005: 65ff.), no longer possesses primacy and the communicative disciplinary regulatory structure of the lessons is not yet established. This leads to a highly unstructured situation characterized by simultaneous coexistence and correlation. Here, actionist practices can take place; practices of experimenting with different gender roles, "for fun", spontaneously and without significant risk of ridicule.

Secondly, these actionist practices exhibit a distance-taking from the institutional order with its symbolic and spatial-territorial elements, 'role distance' in Erving Goffman's (1961) sense. The simultaneous execution of the various communicative rituals of getting ready for lessons – for instance hanging up outer clothing on the coat racks and taking one's seat – is dealt with conjunctively. So Roswitha uses the desk top – whose function, in the context of lessons, is for writing on – as the instrument of her physical encroachment on Robert; and by doing so, she simultaneously underlines symbolically that his taking of his seat is a mere compulsory act. Robert, on the other hand, uses the desk and chairs as an assault course and thereby disregards not only their institutional function but also reverses the institutional spatial arrangement of the room – and, on top of that, this is accompanied by an injury to the inner-territorial boundaries of the seating plan. At the same time the institutional ritual order with its micro-ritual sequences is temporarily suspended.

Such actionist practices occur in the social situation of the transition from the habitual structures of the conjunctive experiential space of the peer group, to the institutional communicative sociality of class, whose regulatory structure is distinguished by standardisation and a reduction of bodily-sensual activities. A considerable number of aspects – which Turner has described as *liminal* – unfold here. These include "kicking over of the traces" and experimentation, likewise the game with the symbolic elements of familiarity or the reversal of social positions. Fun, sensuality and pleasure, spontaneity, rhythm and movement are typical performative forms of conjunctive ritualisation in the liminality, ritualisations, that are in a tension with the simultaneously observable rituals of getting ready for lessons.

On the Multidimensionality of Documentary Video-Interpretation

The empirical example demonstrates the multidimensional quality of social reality, a quality which can be grasped by documentary video-interpretation through the dimensions of time and space and the related sequential and simultaneous structure.

As already mentioned, two different social dimensions can be distinguished: on the one hand the *communicative collectivity*, and on the other hand the *conjunctive experiential space*. The communicative context of meaning, as we find it in institutional fields, produces a role-like mode of acting and a generalized, theoretical knowledge, that is abstracted from the different existential backgrounds and milieu-specific experi-

ences of the actors (see Mannheim 1980: 287ff.). In the video clip the communicative dimension is recognizable, where the interactors refer, by their actions, to the institutional order of the school and to their role as schoolchildren.

In contrast to this, conjunctive experiential spaces (e. g., the milieu of a specific peer group) are rooted in biographical commonalities; i. e. in commonalities of socialization and of milieu-specific social experiences. In these experiential spaces we find an immediate, intuitive understanding and habitual concordance among those who belong to them. The everyday conjunctive practices are structured by collective orientation, which is structured by atheoretical, experience-based practical knowledge (see Mannheim 1980, Bohnsack & Nohl 2003, 369f.).

Furthermore, the conjunctive context of meaning itself can be differentiated. Thus it is characterized by an interlocking of different conjunctive dimensions, of different conjunctive experiential spaces. For example in the video clip, the gender dimension is at the centre of the actionist practices. However, following the comparative analysis of interactions of other children in the same class, this dimension is interconnected with the dimension of adolescence. Thus the “little ones”, as the bigger children call them, do not show such actionist practices. Rather, they are focused on performing rituals of getting ready for lessons.

Methodologically, the documentary interpretation is founded in an analytic attitude, that turns from the “What” to the “How” of sociocultural phenomenon, and thereby to their performative processes. The What, the immanent meaning of the actions and interactions, is analysed in the first stage of interpretation: the *formulating interpretation* (Bohnsack 2003a). Therefore, the elements of conduct and interaction are described in detail without imputation of motives of acting (i.e. “In-order-to-motives”, Schütz 1971: 22ff.). Besides, this stage of interpretation corresponds to the *pre-iconographic description*, the first level of iconology, a method of interpreting works of art, developed by the art-historian Erwin Panofsky (1997: 35f.). The simultaneity is taken into account by describing the physical-spatial organisation and the scenic arrangement of the observed interaction and by analysing simultaneous interactions.³ The explication of the observed phenomenon has yet to come to such a level of detail that the coordination of the elements of conduct and their embedding in long-term interactive processes, as well as in the social situation itself, don't get lost from sight.

The next stage of interpretation, the *reflecting interpretation*, aims at the conjunctive and communicative dimensions of documentary meaning (Bohnsack 2003a), that is the How of the described social situation, interaction and, if for them relevant, the actions of individuals. Thereby, also the way of handling objects, stylistic props and territories, in other words their symbolism and meaning, is of interest.

A central component of reflecting interpretation is *comparative analysis*, which aims to find homologies as well as contrasts in regard to the conjunctive and communicative

³ Due to limitations of space I cannot go into the interactions of other children that took place simultaneously in the video example (see: Wagner-Willi 2005: 300ff.).

dimensions of the observed interactions, according to the principle of *contrast in similarity* (“Kontrast in der Gemeinsamkeit”, Bohnsack 2003a: 37). The systematic comparison is effected with regard to the level of the social situation, e. g. in my study the case-external comparison of the video footage of the phase of transition with video footages of lessons. Furthermore the comparison is also made in respect of the level of interaction inside a specific social situation, e. g. the case-internal comparison of video clips of interactions inside the video footage of the phase of transition. In this way it is possible to identify the communicative dimension (of the institution) as well as the different dimensions of conjunctive experiential spaces – peer group culture, gender, adolescence, etc.

As with documentary interpretation in general, video-interpretation, too, pays attention to the dramaturgy and interrelation of the actors by analysing the *organisation* of interaction. The terms of the explication refer to the verbal, bodily, spatial and material aspects, thus to the simultaneous structure of the organisation of interaction. Often its complexity increases through the sequential structure of dramaturgy: e. g. coexistent interactions can increasingly interrelate; individuals can be involved in ongoing interactions etc. Especially dramaturgic climaxes indicate a high intensity of simultaneity. In the case of discourses, Bohnsack (2003a: 86) calls them *focussing metaphors*, in the case of scenic courses of action, Iris Nentwig-Gesemann (2002: 47) terms them *focussing acts*. Focussing acts afford a particular insight into the conjunctive dimensions of habitualized practices.

In conclusion, video analysis offers a way into *multidimensionality*, i.e., in the first place into the multidimensional sequential and simultaneous structure of the micro-processual interactive praxis of social situations. Documentary video-interpretation is focused on analysing both the multidimensional *performative* structure of the observed interactions as well as their different *conjunctive* and *communicative* dimensions.

The methods of text interpretation have achieved a high degree of complexity – up to now at the cost of *neglecting* the simultaneous structure of social situations. A methodologically reflected video-interpretation has the potential to *close* this gap while maintaining the complexity of text-based qualitative research.

References

- Bohnsack, R., 2003a: *Rekonstruktive Sozialforschung. Einführung in qualitative Methoden*. (5th edition) Opladen
- Bohnsack, R., 2003b: Qualitative Methoden der Bildinterpretation. *Zeitschrift für Erziehungswissenschaft*, 6, 2, 239-256
- Bohnsack, R., 2004: Rituale des Aktionismus bei Jugendlichen. In: C. Wulf and J. Zirfas (eds.) *Zur Innovation pädagogischer Rituale. Jugend, Geschlecht und Schule*. (Zeitschrift für Erziehungswissenschaft, 2. Beiheft 2004). Wiesbaden, 81-90
- Bohnsack, R. & Nohl, A.-M., 2003: Youth culture as practical innovation. Turkish German youth, ‘time out’ and the actionisms of breakdance. *European Journal of cultural studies* 6, 3, 366-385

- Brandt, B., Krummheuer, G. and Naujok, N., 2001: Zur Methodologie kontextbezogener Theoriebildung im Rahmen von interpretativer Grundschulforschung. In: S. von Aufschneider & M. Welzel (eds.) *Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung*. Münster, 17-40
- Garz, D. & Kraimer, K., (eds.) 1994: *Die Welt als Text. Theorie, Kritik und Praxis der objektiven Hermeneutik*. Frankfurt a. M.
- Göhlich, M. & Wagner-Willi, M., 2001: Rituelle Übergänge der Kinder im Schulalltag. Zwischen Peergroups und Unterrichtsgemeinschaft. In: C. Wulf et al.: *Das Soziale als Ritual: Zur performativen Bildung von Gemeinschaften*. Opladen, 119-204
- Goffman, E., 1961: *Encounters. Two studies in the sociology of interaction*. Indianapolis
- Goffman, E., 1981: *Forms of Talk*. Philadelphia
- Huhn, N., Dittrich, G., Dörfler, M. & Schneider, K., 2000: Videografieren als Beobachtungsmethode in der Sozialforschung am Beispiel eines Feldforschungsprojekts zum Konfliktverhalten von Kindern. In: F. Heinzl (ed.) *Methoden der Kindheitsforschung. Ein Überblick über Forschungszugänge zur kindlichen Perspektive*. Weinheim, München, 185-202
- Imdahl, M., 1996: *Gioto – Arenafresken. Ikonographie – Ikonologie – Ikonik*. München
- Knoblauch, H., 2000: Workplace Studies und Video. Zur Entwicklung der visuellen Ethnographie von Technologie und Arbeit. In: I. Götz and A. Wittel (eds.) *Arbeitskulturen im Umbruch. Zur Ethnographie von Arbeit und Organisation*. Münster, 159-173
- Leber, M. and Oevermann, U., 1994: Möglichkeiten der Therapieverlaufs-Analyse in der Objektiven Hermeneutik. Eine exemplarische Analyse der ersten Minuten einer Fokaltherapie aus der Ulmer Textbank ("Der Student"). In: D. Garz and K. Kraimer (eds.) *Die Welt als Text. Theorie, Kritik und Praxis der objektiven Hermeneutik*. Frankfurt a. M., 383-427
- Mannheim, K., 1980: *Strukturen des Denkens*. Frankfurt a. M. (orig. 1922-1925, unpublished manuscripts)
- Nentwig-Gesemann, I., 2002: Gruppendiskussionen mit Kindern. Die dokumentarische Interpretation von Spielpraxis und Diskursorganisation. *Zeitschrift für Qualitative Bildungs-, Beratungs- und Sozialforschung*, 1, 41-63
- Panofsky, E., 1997: *Studien zur Ikonologie. Humanistische Themen in der Kunst der Renaissance*. Köln
- Sacks, H., Schegloff, E. A. and Jefferson, G., 1978: A Simplest Systematics for the Organization of Turn Taking for Conversation. In: J. Schenkein (ed.) *Studies in the Organization of Conversational Interaction*. New York, 7-55
- Schütz, A., 1971: *Das Problem der sozialen Wirklichkeit. Gesammelte Aufsätze*, vol. 1. Den Haag
- Schütze, F., 1983: Biographieforschung und narratives Interview. *Neue Praxis*, 13, 3, 283-293
- Schütze, F., 1987: *Das narrative Interview in Interaktionsfeldstudien I*. Studienbrief der Fernuniversität Hagen. Hagen
- Stadler, H., Benke, G. and Duit, R., 2001: Gemeinsam oder getrennt? Eine Videostudie zum Verhalten von Mädchen und Buben bei Gruppenarbeiten im Physikunterricht. In: S. von Aufschneider & M. Welzel (eds.) *Nutzung von Videodaten zur Untersuchung von Lehr-Lern-Prozessen. Aktuelle Methoden empirischer pädagogischer Forschung*. Münster, 203-218
- Thorne, B., 1993: *Gender Play. Girls and Boys in School*. Buckingham
- Turner, V., 1969: *The Ritual Process. Structure and Anti-Structure*. New York
- Wagner-Willi, M., 2001: Videoanalysen des Schulalltags. Die dokumentarische Interpretation schulischer Übergangsrituale. In: R. Bohnsack & I. Nentwig-Gesemann and A.-M. Nohl. (eds.) *Die dokumentarische Methode und ihre Forschungspraxis*. Opladen, 121-140
- Wagner-Willi, M., 2005: *Kinder-Rituale zwischen Vorder- und Hinterbühne. Der Übergang von der Pause zum Unterricht*. Wiesbaden

- Wulf, C., 1998: Mimesis in Gesten und Ritualen. *Paragrana. Internationale Zeitschrift für Historische Anthropologie*, 7, 1, 241-263
- Wulf, C., Göhlich, M. & Zirfas, J. (eds.), 2001: *Grundlagen des Performativen. Eine Einführung in die Zusammenhänge von Sprache, Macht und Handeln*. Weinheim, München
- Wulf, C., Althans, B., Audehm, K., Bausch, C., Göhlich, M., Sting, S., Tervooren, A., Wagner-Willi, M. and Zirfas, J., 2001: *Das Soziale als Ritual: Zur performativen Bildung von Gemeinschaften*. Opladen