

# Studying social organisation with video

Stuart Reeves, University of Nottingham, 8<sup>th</sup> July 2011

This report was developed from an invited talk I gave for the Research Squared Conference 2011 on Data Analysis held at the University of Loughborough. It examines the role of video within qualitative empirical studies of social conduct and organisation. I focus particularly on the use of video in ethnographies of technology use (mainly because this is my area of interest), and attempt to communicate why video is used to study social interaction. The report provides a broad overview of a particular kind of video use found in various different ethnographic traditions, and particularly focuses on instances within the fields of human-computer interaction, and computer supported cooperative work (e.g., workplace ethnographies of coordination and collaboration around technological environments). I'll look at how video is used, practically, and what techniques and procedures are often employed. We will also cover briefly some underlying commitments to ethnomethodological or conversation analytic approaches that pervade studies of this kind. Finally I'll look at what kind of findings you get from using it by covering a number of well-known studies which used video extensively, and consider what kind things you should be aware of when using video, such as various practical matters, caveats, and so on.

An essential reference is Heath, Hindmarsh and Luff's (2010) book, *Video in Qualitative Research*, which covers much of this material in far greater detail (and with more sophistication).

## **Video-based, interactionist analyses**

There are many different terms used when researchers are engaging in studying social conduct and employing video recordings to do so, and there are many different underlying sociological, theoretical or philosophical commitments that they approach the data with. This report does not pick these apart much, so for now it glosses them broadly as being "video-based, interactionist analyses". Essentially they are about understanding how social order is organized.

<b>"video-based, interactionist analyses"</b>	
	Interactionist commitment
Key resource of study	Meaning is situated
Exhibiting analysis	Order emerges from interaction

So, they use video as a primary resource that is itself studied, often very intensively. Crucially, the recordings themselves are also often used as a way of exhibiting analytic findings to academic audiences. These analyses commonly have interactionist commitments of some form or another, in that they generally identify that meaning and shared understanding as something that emerges ongoingly from the situated interactions of social actors in the setting under study. In this view, order emerges from this interaction, and presents something of a 'bottom up' approach.

Various different disciplines employ video-based interactionist analysis as an important resource in driving analytic work and conceptual development, including disciplines such as anthropology, sociology, or applied linguistics. This report is written from a background of video analytic work as found in computer-supported cooperative work and human-computer interaction. These fields have been interested in using video to understand social organisation with and around technology for a few decades. As such, the way this report approaches this topic is biased towards the traditions and norms found in these fields and the account is therefore limited, rather than comprehensive.

## **Ethnography**

We need to understand the use of video in terms of ethnographic practice. Video-based interactionist analyses across the disciplines mentioned above have typically been performed as part of an ethnography. In turn, ethnographic work has increasingly used audiovisual resources as 'data'.

Ethnography employs participant-observation as a key technique for developing an deep understanding of the member's perspective of the setting under study. As Malinowski describes, the ethnographer's aim is "to grasp the native's point of view, his relation to life, to realize his vision of his world". In this way we can see immediately how video recording can be employed as natural extension of this ethnographic practice.

Video is rarely (or justifiably) collected by itself, and instead is part of an 'instrument' assembly used in participant observation work. Video recording is typically done as part and parcel of collecting together various materials to support the ethnographic analysis and the ethnographic report. It is another resource amongst observational notes, source documents / or other artifacts collected from the field (e.g., office paperwork), collections of log files or computer data files, interviews, etc.

Using these resources the ethnographer aims to arrive at a rich and detailed description of the setting and thus the social organisation of that setting. The richness of video can clearly contribute to that understanding greatly.

## **Using video as an ethnographic resource**

Video enables us to more easily capture the use of technology 'in the wild', in non-experimental and uncontrolled settings. These are kinds of settings that often need highly mobile observation work.

As an ethnographic resource video supports massive replayability. It lets us examine happenings in detail, and can capture moments that couldn't really be seen in-the-field. It also provides an increasingly rich quality of media, as seen with the advent of high definition video.

However, the caveats that come with using any ethnographic data also apply to video. Video is only an aid to understanding a particular social setting – findings don't come *from* the video or any other data; findings instead come from comprehensive understanding of the setting, which may be *assisted* by ethnographic data.

### **Ethnomethodology & conversation analysis**

Next we'll briefly look at the background to much of this video analysis, and the orientations which have informed the way in which many of these studies have been conducted. In particular we are concerned with ethnomethodology and conversation analysis, since both have strongly influenced the way in which video has been used within ethnographic work in CSCW and HCI.

Broadly, ethnomethodology and conversation analysis are interactionist perspectives on social organisation. Although there are distinctions between ethnomethodology and conversation analysis, it's beyond the scope of this document, so I'll uneasily gloss them together as one. (Examples in this report come from both perspectives.)

Frequently we see in HCI and CSCW literatures the term "ethnomethodologically-informed ethnography", which is a kind of signifier for a certain brand of work which presents a certain flavour of findings. Often this involves the use of video recordings both during analytic work and subsequently for presentation to an audience.

What sort of things are we orienting to when we do this kind of ethnography and using video in this way? Primarily these approaches are concerned with the phenomena or 'problem' of social order, i.e., examining how shared understanding achieved and accomplished between members of society. Accordingly, it asks questions such as: What does a member of a particular social setting need to know to be competent in that setting? Why do they do things in this particular way? How is membership organised and coordinated, and how are members inducted into their membership of a particular setting?

Essentially this is, then, about studying the methods employed by members in accomplishing social order. As with ethnographic commitments in general, we approach our study with a deep concern for what the member's perspective is, rather than a theoretical understanding which we then apply to the setting. In this sense (remembering caveats) it is 'data driven' (although such a characterisation is flawed).

In these kinds of studies we are interested in the ignored, taken-for-granted, background features of everyday life. For instance, something as mundane and everyday as queuing is actually an accomplishment that is achieved by the practical actions of members of the setting, in spite of it appearing as though 'nothing is being done'. As Garfinkel reminds us, our interest here is in performing "an organizational study of a member's knowledge of his ordinary affairs, of his own organized enterprises, where that knowledge is treated by us as part of the same setting that it also makes orderable" (Garfinkel, 1967). Often it is these very features of the setting which are important for understanding how technology is used and how interactive technologies may be designed to fit with practice (although obviously there is a relationship between the two which can be tricky).

Using video can greatly assist us as we study the work of interaction in these 'mundane' settings, with this particular orientation to that setting. We can pick apart in great detail how conversation and bodily interaction features in social settings. So, for example, we might examine turn-taking (in conversation) as an interactional device, we might explore how participants construct a sentence so as to be recognisable to others as a 'turn', and how they design their utterances for a particular recipient or recipients.



Above is an example drawn from an augmented reality exhibit in a gallery installation (Reeves, 2011). Here we can see two people (Sally and Tom) collaborating around this technology. The exhibit itself is designed to be like a seaside telescope, which one peer's through the eyepiece of and directs around the space in order to digitally augment physical objects in the room.

The members of this particular social setting are taking turns in using the telescope device, negotiating with one another using their bodily conduct and talk in order to be able to see the same content. The participants in this space build the physical features of the telescope into these bodily interactions and conversational turn taking, so using movements of the device, exchanges of space and so on in order to collaboratively work out handing over the telescope and yet retain and share the same perspective and alignment on the digital content.

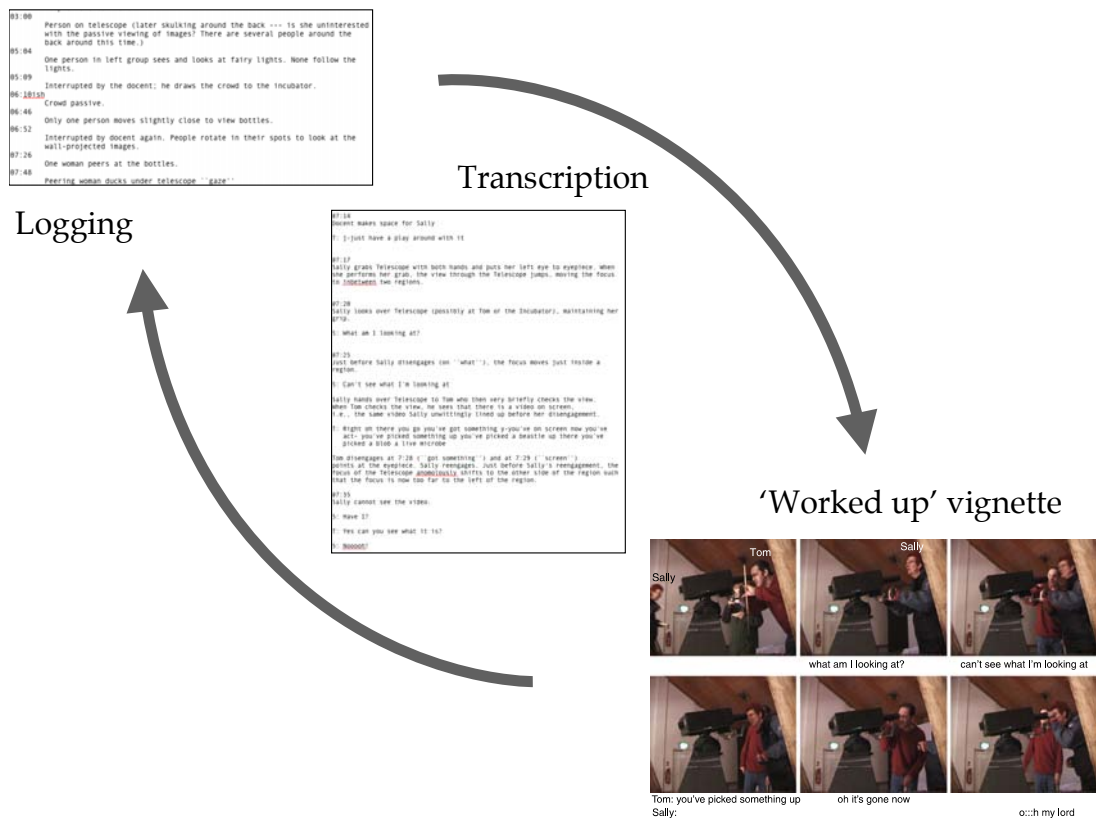
By repeatedly examining instances like this we can come to understand how people practically, really use interactive technologies.

### **The practicalities of using video**

What are the procedures and practicalities of actually doing a study using video? In the case of the interactive augmented reality gallery exhibit shown above, it was possible to leave multiple fixed cameras around the space (as well as 'hang around' and do observation work) for extended periods of time. However, other settings we might study are very short temporally and only exist for a short time, such as studies of ticketed events over a couple of days, performances that happen for a few hours, and so on. Video firstly allows greater access to the latter kind of setting as well as permitting us to capture fleeting moments in both kinds of settings that would have passed participant observation by, or taken a long time to fully develop an understanding for (or only happened once, but were significant).

In order to study settings we might use a number of strategies using various combinations of ambient or mobile video cameras. Generally we can divide up the forms of recording into four different types: we might focus on a particular space that interests us; we could instead focus on an activity that members of the setting commonly engage in within a particular space or series of spaces; we might focus on a particular person, so this could involve shadowing them as they perform their duties; or finally we might try to capture how an object passes through a space or series of activities. There are other ways to capture settings with video recordings, but this gives some notion of just a few of the different techniques.

Once we have collected a corpus of video recordings we need to study them as part of our ethnographic work. In terms of video-based interactionist analyses, there is often a particular process that analysts go through. It should be noted that this is not a 'method' for generating findings. Rather, it is a characterisation of a particular kind of work that is commonly done with video recordings, although the techniques that are employed can vary wildly. As such, this description is obviously partial.



With this in mind, we can consider how repeated review of video recordings begins to help with the analytic work of studying the video data. Often the first stage of this process involves conducting a review of the entire data set, which can be many hours in length, and then logging broadly what the content of the recording is in order to provide an overview of the data set. By doing logging we can get a broad map of the data set, including timestamps and 'what happened' (see figure above). A second part of this process may be then collating segments of interest, such as repeated occurrences of certain activities we are interested in investigating, or repeated occurrences of a particular conversation or even unusual incidents that occur once or twice,

e.g., when something fails. After this, a third stage could involve drilling down into these collections and conducting a focused analysis on them, involving considerably detailed transcription work in order to generate increasingly rich descriptions of the action. Finally, certain vignettes or sequences maybe worked up as part of a publication or other way of exhibiting certain phenomena of social order that is of relevance for the analytic points that are being made.

Again, it is worth noting that this is not a generalised method, but rather is just an account of what kind of things might normally be done. This lack of method or definite procedure for extracting findings is due to the requirements of each setting a peculiar to that setting. So, we don't apply video collection *to* a setting but rather understanding a particular setting may be made *more accessible* via video. As part of this we are using video to support our work in unpacking social organisation in great detail, using video evidence to work through our questions of how things are organised, and just how social order is accomplished physically and verbally.

Furthermore, the process should be seen as iterative but can also reverse back on itself (see figure above), and may involve returning to, say, the logging in order to pick out further incidents having conducted transcription of other events. As such, this process is about developing great familiarity with the data set.

## Transcription

Transcription is a key technique in this process of logging, and repeated review of video data. Below is an example from a paper by Charles Goodwin, and it helps illustrate how detailed transcripts can be.

1 Ray: Doctor Wesley?  
 (0.7) ((Ann turns and walks toward Ray))  
 2 Ann: EHHH HEHH ((Cough))  
 4 Yes Mister Jones.  
 5 Ray: I was gonna see:  
 6 Ann: °Eh heh huh huh  
 7 °eh heh-huh huh  
 8 Ray: Uh::m,  
 9 Ann: Ha huh HHHuh

10 Ray: I think I finally found **this** feature  
 11 (0.8) Cause I: hit the **nail**.

Here they transcribe elements such as conversational turn-taking (1), interruptions between conversational participants (2), pauses in talk (3), and volume (4). We also find moments in talk coupled with images tied to

particular moments in the action that have been taken from the video recording (5). They also annotate physical movements through arrows and enlarge key details (6).

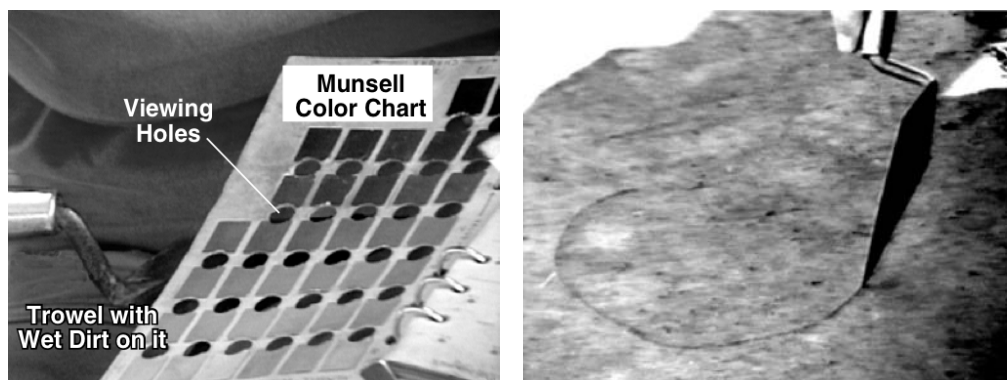
However, we must avoid fetishising transcripts – in the end, they are only *renderings* of what happened (see below).

### **Professional vision**

We can look at a few well-known examples to see more of how this process develops analytically. Charles Goodwin's studies of what he terms "professional vision" (1994) extensively employ video recordings in order to study (amongst other things) the work of archaeologists in the field. He was able to capture the mundane details and practical matters of what archaeologists do in their daily work, and, using video, could repeatedly view those mundane details and develop a rich understanding for how archaeologists visually interpret the terrain of their working environment.

This work unpicks how expert 'seeing' is done on a moment-by-moment basis, within interaction, and how archaeologists can transform a patch of 'dirt' into evidence of an ancient wooden post. They do this through professional, learned seeing practices.

Goodwin's videos document the use of tools such as the Munsell colour chart in order to delineate tiny graduations in dirt colour, the bodily interactions of archaeologists in, say, physically highlighting boundaries of these graduations for one another (see figure below).



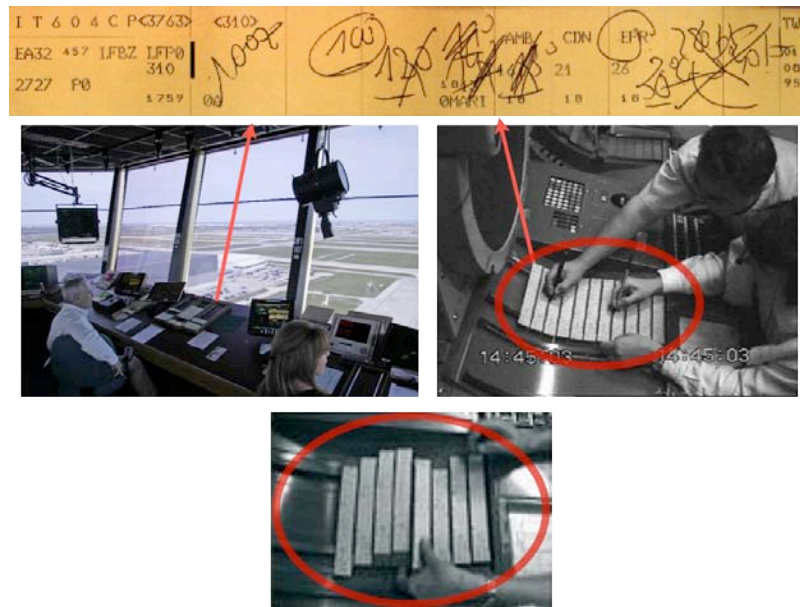
So, these findings are about how the archaeologists organise their expert way of seeing dirt. Being able to do this has in part come from being able to capture with video the detail of the actions of archaeologists' work. Clearly, the mobility of video recording using handheld cameras supports this process.

### **Ethnographic studies of control rooms**

Bentley et al. (1992) studied how air traffic controllers use paper 'flight strips' as a vital part of their work in managing airspace. In the study there is a team of workers who are generally responsible for one sector of airspace. They use radar and communications equipment to manage that space. A flight strip represents a plane in the air through a particular sector, and shows information like call sign of the aircraft, speed, altitude, direction and so on. These strips are put together on a board (see images below).



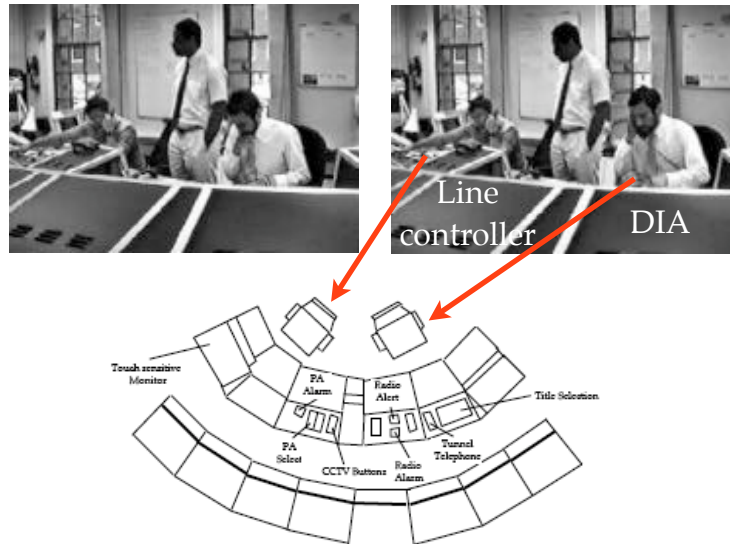
The strips are the focus of the activities of the controllers and how they manage the airspace even though everything is available digitally. Controllers manage the airspace by literally physically managing the strips. Strips can be annotated, folded, moved up and down racks. Their configuration physically – i.e., where and how the strips are organised spatially – indicates things like the status of the flights, or whether someone wants attention drawn to them by pulling them out of alignment.



In studying this setting, coupled with an in-depth competent understanding of the work practices of the controllers, researchers used video recordings extensively in order to monitor certain spaces and the transition of objects, such as the paper flight strips and the rack of multiple strips. This was achieved in part by being able to record large volumes of video focussed on a particular object, thus enabling them to exhaustively compare and contrast how these strips were used.

A further example is a well-known video analytic study by Heath and Luff (1992), in which they observed London underground control rooms. In this study, a Line Controller coordinates the day to day running of a given underground train line. The Line Controller works with a Divisional Information Assistant (DIA), who provides information to passengers through a PA system; he also communicates with station managers (see figure below).





Heath and Luff's study unpacks the interaction between members of staff in the control room. Their study uncovered the ways in which mutual awareness between staff in the control room, drivers and other remote staff was achieved and managed. They looked in particular at the interactions of the physically local staff. For instance, when there was a problem on one of the lines that was reported to the control room, local staff overheard, and started action that took into account their overhearings. They uncovered how staff surreptitiously monitored each other and acted on as well as transformed relevant information. Similarly, staff made certain relevant activities visible for one another to support this practice.

These practices enabled the staff to work together smoothly through organisationally interleaving their activities.

Again, understanding this often highly subtle coordination was enabled through extensive video recording.

### **Attendant issues**

There are some attendant issues which apply to other ethnographic forms of data as well as video, but are nevertheless important to note.

Firstly, the use of video, like any ethnographic work, can impact upon the participants. Often this impact is overstated, particularly for, say, workplaces in which after some 'bedding in' period, the job at hand is always more important. Managing this issue of impact is important, however, and part of this is about developing familiarity with those you are studying, making them comfortable with what you are doing, i.e., recording them, explaining what the video will be used for and attempting to underline that video recording is not about evaluating performance.

Alternatively in my own practice, this issue has been mitigated to some extent by studying public events of different kinds. In these situations there have often been existing expectations of filming or capture of some kind (e.g., theme parks, exhibitions and art performances – see Reeves (2011)).

Secondly, there is a more general issue of 'renderings' we must concern ourselves with. We have to keep reminding ourselves that video offers

limited perspectives on the action we are studying, whether this is in camera framing, or in choices about where to place cameras at all. Video should instead be seen itself as a *form of transcription*.

Thirdly, there are challenges in being rigorous in investigating a particular social setting, and how generic findings can be. There are no easy solutions to this, however becoming a competent member of the settings under study is an important way of dealing with this, as well as the way we treat our data. So, we might relentlessly collect together repeated instances of the same event, e.g., a particular activity being performed, such as use of the flight strips or instances of when people handed over the exhibit, or where actions between control room staff appear to be linked.

At the same time we must keep in mind that 'one off' events such as breakdowns and moments of failure can also prove crucial to our understanding and analytic work. Typically, in providing a moment of 'breaching' (Garfinkel, 1967; Crabtree, 2004), they reveal a lot about how settings are organised socially and how they are worked out practically, moment-by-moment.

Finally there is a tricky relationship between producing ethnographic descriptions of a setting and how that is then used to inform technology design. Video analysis work within ethnography, and indeed ethnography in general, cannot just be a 'servant' of design.

## References

- R. Bentley, J. A. Hughes, D. Randall, T. Rodden, P. Sawyer, D. Shapiro, and I. Sommerville (1992) Ethnographically-informed systems design for air traffic control. In *Proc. CSCW '92*. ACM, New York, NY, USA, 123-129.
- C. Heath and P. K. Luff (1992) Collaboration and control: Crisis management and multimedia technology in London Underground line control rooms. *Journal of Computer Supported Cooperative Work*, 1(1-2):69-94.
- H. Garfinkel (1967) *Studies in Ethnomethodology*. Prentice-Hall Press.
- S. Reeves (2011) *Designing interfaces in public settings: Understanding the role of the spectator in Human-Computer Interaction*. Springer.
- C. Goodwin (1994) Professional vision. *American Anthropologist*, 96(3):606-633.
- A. Crabtree (2004) Design in the absence of practice: breaching experiments. In *Proc. DIS '04*, pages 59-68, New York, NY, USA. ACM.
- C. Heath, J. Hindmarsh, and P. Luff (2010) *Video in Qualitative Research*. SAGE.