

Some 'technical challenges' of video analysis: social actions, objects, material realities and the problems of perspective

Qualitative Research

12(3) 255–279

© The Author(s) 2012

Reprints and permission: sagepub.

co.uk/journalsPermissions.nav

DOI: 10.1177/1468794112436655

qrj.sagepub.com



Paul Luff

King's College London, UK

Christian Heath

King's College London, UK

Abstract

Unlike the wide-ranging methodological debates surrounding the accomplishment and analysis of interviews, fieldwork and focus groups, the discussions concerning the use of video data tend to focus on a few frequently rehearsed issues. In this article we wish to broaden the consideration of methodological concerns related to video. We address the problems faced when collecting data, particularly on how to select the framing for the recordings. We discuss the problems faced by researchers and how these have been addressed, revealing how a conventional solution has emerged that facilitates a particular kind of 'multi-modal' analysis. We then suggest some limitations of this framing and describe a number of recent approaches to recording video data that seek to overcome these constraints. While providing opportunities for very distinctive kinds of analyses, adopting these solutions places very particular demands on how data are collected, how research activities are conventionally undertaken, and perhaps more importantly, the nature of the analysis that is made possible. Although seeming to be a practical and technical consideration about recording data, selecting a camera angle uncovers methodological concerns that reveal the distinctive demands that video places on researchers concerned with the detailed analysis of naturally occurring social interaction.

Keywords

interaction, video, workplace studies

Corresponding author:

Paul Luff, King's College London, Department of Management King's College London, Franklin-Wilkins Building 150 Stamford Street, London SE1 9NH, SE1 9NH, UK.

Email: paul.luff@kcl.ac.uk

Introduction

Despite the growing interest in the use of video to support qualitative naturalistic research, save for a few notable exceptions (e.g. Knoblauch et al., 2006; Pink, 2006) there are still relatively few discussions of the issues that arise when it is used. Indeed, unlike methodological debates surrounding the accomplishment and use of interviews, fieldwork and even focus groups, the discussion concerning video is still constrained to a small number of issues. These principally focus on the difficulties associated with gaining access to record, the ethics of collecting video materials of everyday conduct and the potential effect of recording on 'naturally occurring' activities. To ameliorate the problems identified, generic solutions are typically proposed, guidance that is potentially useful for a wide range of qualitative researchers. Despite the importance of these concerns, video raises a far wider range of issues than those typically discussed. Indeed, its use presents challenges at almost every stage of the research process, whether these are: choosing where, what and when to film; how to categorize and transcribe the data that are collected; how to select fragments to analyse from a corpus; how to develop an analysis that resonates with the collected materials and how to present analyses to audiences or in conventional research publications. Although such difficulties are faced by qualitative researchers managing other kinds of data, the problems associated with video do seem distinctive, not only because of the rich nature of the material collected, but also because the potential ways of resolving them seem to vary according to the analytic orientation adopted by the researcher. In this article, we wish to discuss one such problem faced when using video. This occurs at the earliest stage of a research project and it might seem straightforward to resolve. This is the problem of selecting the angle and framing of the recordings. The choices made when recording can facilitate or undermine subsequent analysis. We also suggest that these choices raise a number of critical methodological issues and how these are addressed cannot be divorced from the ways the collected materials are subsequently analysed.

In this article we focus on research where video recordings are drawn upon as the principal materials for an analysis of naturalistic conduct. This contrasts to the use of video collected to develop into different kinds of documentary resource (e.g. Pink, 2004, 2006), to provide a richer recording of an interview or gathered to provide illustrations for an associated analysis. The video data is also later subjected to detailed analysis, typically of the social interaction between the participants recorded. Such naturalistic recordings have been drawn upon by researchers in sociolinguistics (e.g. Kendon, 2004), social psychology (e.g. Kress et al., 2005), human geography (e.g. Brown and Laurier, 2005) and sociology, and have informed analysis of workplace activities in a range of diverse settings including healthcare, classrooms, control rooms and design practice (e.g. Murphy, 2005), in public settings such as museums (e.g. Meisner et al., 2007; Vom Lehn et al., 2001) and in domestic environments (e.g. Goodwin, 1981, 2004; Tulbert and Goodwin, 2011). For each, the data collected make available a resource that allows what is often termed 'multi-modal' analysis, the repeated scrutiny at extraordinary levels of detail of how talk and visual conduct in the material environment. In this article we consider how simple choices involved in the collection of video materials can support this kind of study. We draw on a large number of studies that have sought to use video,

many have been reported in the literature but others, particularly where problems have emerged when data are collected, are not. Compared to similar seemingly practical problems raised in both text books and research monographs concerning other data collection methods, the difficulties of recording video and the consequences for subsequent analysis have been little discussed in the literature. We will therefore draw on materials collected in a number of unreported studies in which ourselves, colleagues and students have been involved and where problems and difficulties have emerged. We suggest that a 'conventional' solution has emerged when addressing problems faced when recording naturally occurring video data. However, by adopting this convention certain constraints arise that delimit the kinds of analysis that are possible. We discuss a number of recent approaches to recording data that seek to overcome these constraints, involving the integration of additional recordings and the collection of associated data through conventional fieldwork and interviews. We suggest that while providing opportunities for very distinctive kinds of analyses, adopting these solutions also place very particular demands on the ways those additional materials are collected and integrated, how seemingly conventional research activities are undertaken, and perhaps more importantly, the nature of the analysis they make possible. Although seeming to be a practical and technical consideration about recording data selecting a camera angle uncovers methodological concerns that reveal the distinctive demands that video places on researchers concerned with the detailed analysis of naturally occurring social interaction.

Background: the practical challenges of video analysis and framing

Despite the growing use of video to support qualitative research, to supplement fieldwork for an ethnographic study, to record data from experiments or to provide a way of capturing the responses of interviewees, it still seems to be under-utilised in the social sciences as a means of gathering data concerning everyday social interaction. This is despite innovations over a century ago in adopting video, or more accurately the moving image, to understand naturalistic conduct. Haddon, for example, used film as part of his field work for his studies of Torres Straight Islanders in the late 19th Century (Haddon, 1993). In the early 20th Century Gilbreth filmed everyday work tasks like brick laying (Gilbreth, 1908, 1911, 1973 [1908], 1974 [1909]) to gain an understanding of the organization of routine work. The moving image has also been a longstanding resource in studies of non-verbal behaviour (e.g. Birdwhistell, 1970; Condon and Ogston, 1966; Gesell, 1925; Gesell et al., 1940; Haviland, 1993; Schefflen, 1973). Given these foundations it is perhaps surprising that now, when the technology for capturing very high quality moving images is relatively cheap, video is so little used for the collection of qualitative data in the social sciences. It is hard to argue that participants in settings will be unfamiliar with the capabilities of the technology, will necessarily be resistant to its use or the presence of a camera will drastically transform their conduct. Video cameras are now commonplace in everyday devices such as the mobile phone, in many settings video recording technologies (like CCTV) are ubiquitous, and video is frequently used in training and education. Moreover, the common availability of video editing systems for domestic

purposes means that many of the problems associated with managing large collections of video and manipulating video for research have been ameliorated. What do seem warranted are the problems that arise when attempting to integrate video within certain kinds of analytic frameworks. The quality of the data collected often undermines the development of an analysis; the data collected seeming strangely resistant to analysis. It can then be difficult to codify and categorize video data, to break it apart into fragments that can be subjected to analysis, and transcriptions of visual conduct can seem unwieldy and time-consuming to produce, even for the shortest fragment.

One longstanding problem, reported in different ways by Bateson and Mead (see Hockings, 1995 [1974]), Laurier and Philo (2006) and Macbeth (1999), is that of data collected through the 'roving' camera. Although it is tempting to consider that in following the action the researcher will be able to react to the circumstances occurring in a setting and maybe also record more interesting footage, the material collected in this way is often hard to analyse. Not only do objects and actors move from the frame and at important moments of transition appear out of focus, but often critical aspects of the production of an activity are missing (see discussion in Brand, 1976). For example, in a project examining behaviour in public places, the researchers collected video data recorded in a street in a town shopping centre (see Figure 1).



Figure 1. Filming a street scene with a roving camera

The researcher collecting data first focuses on a group containing a mother and her two children (1.1) and then adjusts the camera angle zooming into the mother. The woman turns sharply and looks to her right (1.2). The researcher then readjusts the camera angle, seeming to follow where the woman is looking, zooming out (1.3) and moving to the left the researcher finds a group of street musicians (1.4). In this fragment something about the behaviour of the musicians has attracted the attention of the children and their mother seems to orient to this shift of attention.

This small clip captures a small moment of collaborative activity in a public setting. The mother seems to be responsive to something her boys, and maybe others in her local vicinity, notice in the environment. However, neither the activity that initiates the looking nor the ways in which the looking emerges with respect to the shifting features of the environment are open to scrutiny. The material can serve to illustrate a claim *that* participants are sensitive to others in the environment and perhaps draw on that conduct to shape their own. However, as it is only possible to speculate on the interrelationships between the activities of the various participants and the environment, *how* that conduct is produced and recognized by participants is not open to scrutiny. The roving camera can be excellent for gathering materials for the purposes of documentary or illustrative

use, but the data collected in this way often undermines the development of detailed analysis of how participants produce and accomplish their actions and activities.

Perhaps the most obvious way to address such problems is to opt for a fixed angle and set this wide enough to capture a larger area so that most participants and features of the environment that could be of potential analytic interest or concern would be visible. In the following project, the researcher interested in public behaviour in and around places in a park adopted such a solution (see Figure 2).



Figure 2. Filming an open air setting with a fixed camera angle

Given such a wide angle, the details of many activities occurring in the setting may not be visible. However, there are technical solutions to the problem. Simple adaptations to the camera lens can help to capture more of the environment, particularly within circumscribed settings, like offices, or control rooms where there may be little option in where to place the camera. When a conventional lens constrains the space that can be viewed, wide angle adaptors can be fitted to a lens so that data can be recorded encapsulating more of the local environment. In a study of an architectural practice (Figure 3, left) a camera with a wide angle lens was placed in a top corner of the open-plan area where most of the design activities took place. This angle captured most of the room and it was possible to record moments where professionals from different specialisms would engage with each other in discussions about ongoing projects. In a study of a control room (Figure 3, right), the wide angle lens allowed the researchers to record both the talk and the visual conduct of staff, in a setting where the positioning of the camera was restricted.



Figure 3. A wide angle lens used in a study of an architectural practice (left) and a control room (right)

Even if it is possible to capture a large scene where some relevant activity happens, the material may not be conducive to analysis. In the study of the architectural practice, the recording revealed that the designers were discussing issues regarding the designs, but neither from the visual materials nor from the audio was it possible to recover what they were discussing. The materials did not provide access to the detail of the participants' visual conduct, particularly to their activities with material artefacts, like the paper plans or the computer systems that were critical to their work. In the control room study the wide-angle lens did provide greater access to the work of the controllers but the wide angle lenses distorted the image dramatically. From the collected data it was difficult to recover the orientation of the participants, both where they were looking at any moment and what objects they might be looking at.

However, even if it is possible to place a camera in a static location that is neither too close to the domain of action or too far away, there can be problems selecting an appropriate framing for recording. As in many control rooms, in the following setting most of the critical activities of personnel seemed to occur on/or around a console. From this set of desks, controllers could manage the power supply to the service, communicate with train drivers and make broadcasts to passengers. As most of the operations of the transport system were coordinated from what seemed to be a circumscribed locale, one framing could be used to capture most of the critical activity in the control room (see Figure 4a).



Figure 4a. Wider angle toward the left of the desk

This shot captured moments when staff collaborated with each other, but what is not available for analysis are the resources the staff use for their work: the screens, documents and keyboards; what the participants say to each other; or the details of how they orient to each other to identify who they are addressing. Again, it is possible to record *that* participants collaborate but not *how*. On the other hand, if the camera angle is too narrow (see e.g. Figure 4b), there is a danger of failing to capture any activities of the participants.



Figure 4b. Close up to the left of the desk

Even in simpler settings it can still be difficult to find an angle with which to record video materials. For example, in a study of Sushi restaurants, a researcher placed the camera on the counter and focused the angle on customers at the end of the bench (Figure 5).



Figure 5. Recordings of customers in a Sushi restaurant

This angle provided a view that gave visual access to how pairs of customers coordinate their selections from the conveyor belt. It also allowed for the capture of talk between pairs of customers, particularly as they discussed the options from the menus. Unfortunately the audible data did not match the visual. The pairs of customers who could be heard were not the same as those that could be seen.

For anyone who has endeavoured to capture video materials these problems should seem familiar and also, perhaps, a little mundane. The use of a static camera is one way of avoiding the problems associated with the use of the roving camera; of following, and often missing 'the action'. Avoiding too wide an angle depends on a better understanding of the activities that are to be recorded and this requires undertaking fieldwork in the setting. However, more familiarity with a setting may not necessarily facilitate identifying a place to start recording. Indeed, this may require the kind of analysis of

participants' activities that video materials make possible. Hence, a convention has emerged for selecting an initial framing for a video recording in everyday settings. This selection may then serve to help identify further shots or angles to record, or may often be the only angle that is used in the subsequent analysis.

The emergence of a conventional solution: the mid-shot

The solution that has emerged and is utilised by a wide range of researchers, including those that draw on video to undertake workplace studies (Goodwin, 1995, 2003; Goodwin and Goodwin, 1987; Heath and Luff, 2000; Mondada, 2007; Whalen, 1994, 1995a, 1995b; Whalen et al., 2002), analysis of activities in public and semi-public settings (Laurier et al., 2001; Vom Lehn et al., 2001) and domestic environments (Goodwin, 1981) can be characterized, in filming, as the medium- (or mid-) shot. More precisely, this is a static shot that has an open angle that typically captures the activities of two to three people. So, for example, in a study of medical practice that was concerned with how doctors used computers within the consultation, the camera was placed in the top corner of the room. It was then focused down on the corner of the desk where the computer was placed, capturing a region that would include both the doctor and the patient when they were seated (see Figure 6a). A similar angle was used in a study of an architects' practice (see Figure 6b). Here, the camera was placed on a tripod across the desk from one particular architect. It captured an angle from his computer to the area where he placed the large paper plans of the buildings he was working on. In a different study of design discussions over paper plans and sketches, the researcher placed the camera on a tripod toward the end of one of the large tables the designers sat at to discuss preliminary ideas (see Figure 6c).



Figure 6. Recordings of a medical consultation (left), architects working on paper and computer (middle) and paper work in a design practice (right)

In each study the camera angle is placed on a tripod and typically captures two to three people in a scene. It is also placed slightly above the eyeline of the participants when they are seated, angled down so that artefacts like paper documents and keyboard are in the shot. Although in each of the studies it was recognized that the participants would not always be in shot when they entered or left the consulting room, this angle does allow materials to be gathered for an analysis of the coordination of the conduct between the participants. With a microphone placed near the participants, the video recording provides materials to support an analysis of how the talk and visual conduct of the participants are interleaved. The camera can be left recording for an entire morning

or afternoon, capturing a whole series of consultations or a number of different discussions. So, in the study of medical consultations, apart from when a patient entered or left the consulting room or when they were being examined on the bed, it provides access to most activities where the computer is used by the doctor or is referred to by the participants. The angle therefore facilitated an investigation of how computer use facilitated or undermined the consultation (Greatbatch et al., 1993, 1995a). In the design and architecture practices, although the angle did record long periods when architects worked alone or when the positions were not occupied, it did allow discussions to be captured between two of three colleagues. This supported analyses of how the details of what was written and drawn on paper were utilised in those discussions, and how these were juxtaposed with images in books and photographs and diagrams presented on computer screens (Luff and Heath, 1993; Luff et al., 2009).

Although when using the stable mid-shot there are often long periods of time when none of the participants are in shot or when one or more are only partly visible, this angle can facilitate the analysis of fragments when two or more participants are engaged in concerted action. It can also capture materials to allow for an analysis of the initiation of activities and how the upcoming completion of an activity can be oriented to by the participants in a setting. These advantages of the stable mid-shot mean that it has been utilised in a wide range of studies of circumscribed activities in work domains as diverse as offices, call centres and control rooms (Suchman, 1997; Whalen, 1995b; Whalen and Vinkhuyzen, 2000; Whalen et al., 2002, 2004). Also, when focused on the area around one exhibit and artefact, the stable mid-shot has been used in studies in public settings like museums and science centres (Heath and Vom Lehn, 2004, 2008; Meisner et al., 2007; Vom Lehn et al., 2001). With the camera being left on the tripod this choice has the advantage of not requiring the researcher to position themselves close to the camera, and so recording can seem less intrusive. The materials gathered using this angle have been fruitful in allowing the production of naturalistic activities to be scrutinized in detail. Through an analysis of the moment-to-moment accomplishment of activities, researchers have revealed how talk and visual conduct, gaze, gesture and bodily comportment are interweaved to produce concerted action, the shifting forms of participation in and around those activities and the different ways technologies, tools and objects (whether they are sophisticated or mundane) feature in everyday conduct.

The angle focuses on small gatherings, typically on participants who are seated, often where only their head and shoulders can be seen, and provides for interactions between those participants to be analysed. Its selection reflects a number of analytic commitments. First, it reflects a concern with the problem of relevance by attempting to draw upon the participants' own perspective; what they are oriented to in the production and accomplishment of action. A stable shot seems to provide some access to the participants' shifting orientation and participation in social actions. It is also possible, to some extent, to recover the orientation to activities of the participants in the environment. Second, the angle reflects a concern with uncovering how actions are accomplished in and through interactions between participants. In particular, by seeking to provide access to two or more participants at the same time, this choice of angle can support the development of sequential analyses of the organization of conduct. Third, through a consideration of the sequential organization of conduct, interaction can be both the topic of the study but also a resource with which to develop analyses. The materials gathered can

facilitate and help make sense of the skilful achievement of activities of participants in many complex domains.

A simple choice of camera angle can serve as an important way to resolve many of the problems faced when starting to record video data in naturalistic settings. It can aid in gathering data, which, being stable, can be systematically reviewed not only by the person recording them but by colleagues. The close circumscribed shot can also facilitate the selection of fragments to commence an analysis. Only focusing on two or three participants can also make the materials more amenable for transcription. Moreover, such a framing provides a way of commencing data collection, allowing the possibility for more sophisticated options to be taken later when the researcher is more familiar with the setting and the phenomena being addressed. Perhaps, for these reasons, using the stable, mid-shot has become the conventional way that audio-visual materials are recorded in workplace studies. The angle seems to allow for the analysis of how social activities emerge and are produced through interactions between participants and the manipulation of objects and artefacts, tools and technologies within the local environment. However, the mid-shot is also restrictive. It only covers one angle and delimits the environment that can be subjected to scrutiny, not only with regard to the number of participants being focused on but also the objects and artefacts that feature in everyday conduct. For complex multi-party activities, like those in classrooms, for different kinds of performances or other forms of behaviour in public, this conventional framing can seem unduly limiting. Even when considering the detailed interaction between participants in focused settings this framing may be too restrictive. The angle can reveal that a document or a computer features in a collaborative activity, but not necessarily how. What is accessible to the participants (e.g. the contents on the page or screen, or the object to which the participants are oriented) is not necessarily available in the recordings. So, even for conventional workplace studies the mid-shot angle, by restricting access to certain features of conduct, can limit the range of actions and phenomena that can be subjected to scrutiny.

Broadening the perspective: multiple views

Perhaps the obvious way of addressing the restrictions of the single, stable mid-shot angle is to gather data from multiple cameras. The advantages of this framing can be maintained by each camera being focused on a mid-shot, but each covering a different angle. So, in a study of a large transport control room the initial data collection used the conventional medium shot angle focused toward the principal participants: a controller and an assistant sitting at a large console (see Figure 7a). The setting was complex. As well as the controllers' desk, the staff responsible for managing the signals sat on a similar console in the same room. Also, in front of each participant there were three or four dedicated screens to display information about the line, operate the telephone or provide CCTV images. Moreover, running along one wall of the room in front of a desk was a large display showing data of the real-time operation of the service. Further data collection included the initial angle but also involved capturing materials from other angles. A second camera was positioned behind the participants (Figure 7b) so the large display could be seen, as well as some of the computer systems used by the participants. A third camera angle focused on the controllers from the side, capturing more computer



Figure 7. Three conventional shots used in a study of control rooms, from in front of the console (left), from behind (middle) and from the side (right).



Figure 8. A mixed image used in a study of control rooms from three different cameras recording at the same time, from in front of the console (top left), from the side of the controllers (bottom left) and from the side of the signallers' desk (bottom right).

screens as well as objects, in particular documents and phones, that were placed in the console (Figure 7c).

The initial angle toward the controller and assistant could support an analysis of how one participant was sensitive to the activities of the other; for example, how an assistant might overhear the controller's phone call with a train driver and then take some relevant next action related to what was in the call (Heath and Luff, 1992). With the second camera angle the recordings might help reveal how controllers might transform what they were doing after noticing an assistant looking at the large display. The third camera angle, from a side view, provided more details of the participants' conduct in relation to the individual screens and one another, but could not capture any details of the large display. Each angle on its own would not facilitate an analysis of how command and control emerged through an interweaving of social interaction (talk and visual conduct) and the technologies available to the participants. Gathering from all cameras at the same time using a field video mixer (see Figure 8) did facilitate such an analysis (Heath and Luff, 1996).

With multiple camera angles it can be possible to recover features of the orientation of the participants. For example, when trying to make sense of something as seemingly simple as when one participant is pointing to an object on the other side of a room, it can be possible to identify what the participant is pointing at, when the gesture is noticed by a colleague and how that colleague transforms their conduct in the light of the actions of others and also particular, often fleeting, features of the environment. However, materials from multiple camera angles can be hard to manage. It can be difficult to analyse images from synchronous sources when presented at the same time. This may be why, when mixed images are available, analysis tends to focus on one of the angles with others used to provide additional resources of concern only when they seem relevant; for example, when it seems important to consider what was available on a screen or where a participant was oriented at a particular moment.

Multiple cameras can also seem invaluable when the environment of interest is less focused and when a conventional framing would seem to undermine analysis of the conduct in the setting. This seems to be particularly the case in settings where there are a large number of participants involved in the principal activity; for example, when audiences participate in an ongoing activity. Consider auctions, for example, where any member of the audience could participate as a bidder. Here, the mid-shot would seem infeasible. A single framing focusing on the auctioneer (see Figure 9a) can reveal features of his or her talk and visual conduct; however, this provides little access to the conduct of the bidders. If a wider angle is used, which includes some of the audience, occasionally it may be possible to identify a bidder, if that bidder happens to be in the shot and the conduct is fairly gross, like the wave of a hand, a sheet of paper or catalogue (see Figure 9b). As well as the conduct of the auctioneer being less visible, this angle precludes access to the more delicate actions of bidders, actions that are typically characterized as being discreet, such as slight nods and glances. The organization of the activity seems to rest on such conduct being unavailable to public scrutiny.



Figure 9. Recording data in auctions. On the left a single image, a close-up of the auctioneer reveals how simple artefacts are utilised in the activity, on the right a wider angle provides some access to the coordination of conduct between the auctioneer and a bidder (on the left of the image).



Figure 10. A mixed image includes one camera angle focusing on just the auctioneer and the other two are of the audience, from the side of the room.

If multiple camera angles are used then these can provide greater access to the conduct of the participants in the environment; for example, taking different views of the audience (see Figure 10).

None of these angles can be considered to be conventional framing of shots and yet through close scrutiny of the recordings of the audience it is not only possible to identify when participants actually make bids or attempts to bid, but also the actions they undertook in preparation to bid, their activities following the auctioneers' attempts to get them to bid again, and the ways in which their participation fluctuated and was transformed with respect to the conduct of the auctioneer. It is possible to develop an analysis of a trajectory of actions, between a principal participant and members of a large audience, through which the activity is accomplished (Heath and Luff, 2007a, 2007b). Broadening the perspective by using multiple cameras provides a means of capturing, to some extent, the fluctuating forms of participation from members in a setting, particularly in complex domains.

Despite providing greater access to a setting and allowing for the detailed analysis of conduct between large numbers of participants or with different kinds of tools and technologies, using many camera angles can present a number of distinct challenges for analysts. It may not be that straightforward to recover an activity from recordings taken from different camera angles. In particular, if the concern of the analysis involves assessing and accounting for the participants' own perspective on the ongoing activities, it can be problematic to determine how the conduct recorded by each camera is seen and understood by the different participants. For example, in the auction materials, a glance or turn toward an

area of the sale room by the auctioneer may be noticeable to a viewer of the close up image, but it is not clear whether it is quite so visible to all bidders in the environment. More importantly, even with two or more different views of the audience, ascertaining how another's conduct is seen by different participants can be problematic. Even if, from the material available, a researcher can be confident *that* a participant can see an activity by another and from their subsequent conduct that they noticed it, it can be unclear *how* they saw it – their particular perspective on the activity at hand. Consider again the case of the auction. A bid might be produced so that it is only noticeable by the auctioneer, or, in a setting like a control room, a gesture may be configured in such a way to not only to be visible to a colleague but to be visible in different ways to different participants. Simply by being two-dimensional the camera image transforms how the production of an activity is seen in a recording from how it was originally produced. In materials where an action is particularly delicate and hard to differentiate, particularly in settings where there are asymmetries between the ways in which conduct is visible and how it is produced, it can be very difficult to assess how an action is shaped with respect to other participants in the setting. Even if the relevant conduct is captured in the recordings, it may be hard to ascertain how that conduct is both produced and understood from all of the perspectives of the particular participants. It is not just problematic to recover the temporal order of actions and events, but also in reassembling how participants see, recognize and produce actions with regard to the activities of others in an environment of action.

Attending to the details: material resources, tools and technologies

While the conventional mid-shot may constrain the view to only a circumscribed area of a setting, for some kinds of analyses it may still not provide enough access to details within that locale. For many activities, work relies on material and digital resources; whether these are documents or computer systems, or more simple tools like, hammers, pens and scissors. Consider the introduction of computer systems over the past 30 years into most workplaces. Video would seem to provide a resource for revealing the details of how work is accomplished with and through these technologies, particularly how they might support or undermine collaboration. It is therefore not surprising that within workplace studies the use and operation of computers has become one focus for analytic enquiry. For example, video recordings of the work in contemporary journalism might allow for a consideration of how journalists in a newsroom not only select 'feeds' from texts made available on news agencies' computer systems, but also transform them in the production of stories and communicate their relevance to colleagues (Heath and Nicholls, 1997). Recordings in electronic dealing rooms can reveal how financial traders monitor the complex information on their computer screens and the details of how they deploy collaborative strategies to trade in stocks and shares (Jirotko et al., 1993). A consideration of how users of systems designed for individuals, either in settings like control rooms or in classrooms, can reveal how they are used within interactions by colleagues (Luff and Heath, 2000; Luff et al., 2008). In each video materials can be drawn upon to reveal the systematic ways in which participants, in each setting, can make what seems private, public; how they may select and make apparent details of their work to colleagues (see Figure 11).



Figure 11. Video recordings captured in a variety of environments where computers are used: journalists drawing upon news feeds to produce stories (top left), financial traders using stock exchange systems in a dealing room (top right), scheduling and communication systems in a transport control room (bottom left) and students discussing statistical data in a student classroom (bottom right).

Video recordings can support the analysis of how such computer-mediated activities are achieved; the conventional mid-shot framing facilitating the analysis. By being stable it can help reveal how one participant initiates an activity, how they might shape their activities in ways that make something that is not visible apparent to a colleague and how these may be transformed by the ongoing activities of that colleague (or colleagues). So, for example, in the study of the use of computers in medical consultations, the angle from the side and slightly above the participants provided materials for an analysis of the coordination between the doctor's typing on the system and the patient's conduct (see Figure 12).



Figure 12. Computer use in medical settings. The angle can reveal when the doctor is typing, some details of the screen and also the patient's orientation to the doctor's conduct.

Using this angle it is possible to have some access to the contents of the screen and the shifting orientation of the participants and hence support an analysis to reveal that patients were not just aware of the doctor's activities on the computer, but were sensitive to particular features of their conduct. For example, a patient might coordinate the initiation of particular kinds of utterance, unsolicited statements, with the doctor's completion of a typing activity, displayed by typing the enter key, particularly when this was struck with force and marked the end of a sequence of actions (Greatbatch et al., 1995b). The conventional stable framing can also capture a trajectory of collaborative activity. In this study it helped reveal how patients also seemed sensitive to the upcoming completion, projecting to speak as the doctor's hand moved toward this key. Even though it was not possible to see details of the screen from the recordings, this kind of analysis was possible because the patients also rarely looked at and read details off the screen. On the other hand, when considering the doctor's conduct (e.g. whether topics for discussion reflected particular items displayed on the computer system) such critical details are not available for scrutiny from this angle. In many cases, such limits of the video record can be less consequential. The constraints for the analyst reflects those of the participant. As in the case of journalists above: colleagues cannot see the details of a co-participants' screen and the resolution to this problem, by making the contents publically available through talk, are open to analytic scrutiny. However, when the details are commonly available to participants (e.g. when details of objects are visible to participants but are not recorded) then the single video recording from the conventional angle can be limiting. There are ways of recovering some of these details, but they are not necessarily straightforward. Some details can be made available for analysis by considering the ecology of the screen, where, for example, changes to the display seem to be occurring. Similarly, if these are visible it can be possible to recover some of the details of the activity by mapping their finger movements across the keyboard (Luff and Heath, 2000). In such cases, where the details of activities rely on a particular artefact for their accomplishment it may be necessary to combine the mid-shot with a close up angle taken from a second camera. In a study of design practice, the researchers were concerned with the collaborative accomplishment of creative work (see Figure 13).

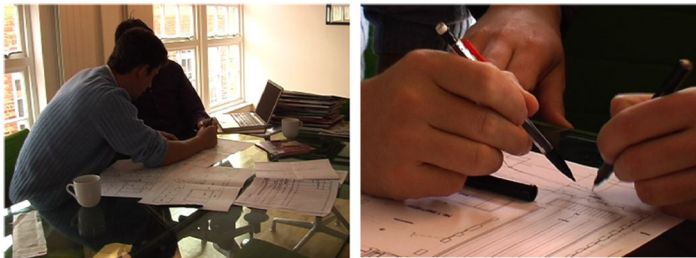


Figure 13. Using multiple cameras to capture the details of artefact use in a design practice. The conventional framing (left) from a camera on the left is augmented by material collected from another camera (right) from the right which captures just the hands and pens pointing and sketching in drawing. In this case the researcher would occasionally readjust the angle of the second camera so it captured the activity which the designers attended to. However, the shot has held stable for most of the time and only the fixed recordings were useful for analysis.

The combination of the two different camera angles facilitated the analysis of the early stages of design, and also how these ideas were made concrete through the collaboration over and around sketches on paper (Luff et al., 2009). When considering conduct with and around material objects, gathering data from a second camera may not always be straightforward. The positioning of the second camera can be problematic, partly because the person manipulating the object in question can obscure the details required. In some settings it can be possible to gather data using more sophisticated methods, involving feeds taken from the monitor, as in studies by Whalen et al. (2002) who have explored how the use of technologies is integrated with the interactions between callers and call takers in a range of call centres. These data, coupled with audio data from associated phone calls, support an analysis of how topics raised by personnel are shaped by the way in which they appeared and ordered on the computer system, but also how the call takers, who had varying levels of expertise, would also transform any guidance on the screen to be more appropriate to the local interactional context of the call (Whalen, 1995b; Whalen and Vinkhuyzen, 2000; Whalen et al., 2002).

However, the use of video recordings to analyse the use of technologies and objects in workplace settings cannot be accomplished by only scrutinizing the recordings. Indeed, the analysis of the video materials that have been collected typically requires considerable knowledge about the activities of the participants. Members of the setting typically bring to bear a range of resources when using a technology or artefact, including expertise and skills associated with the domain, typical problems that occur in the setting, and activities and issues that have arisen in preceding periods of activity. It is therefore necessary to undertake fieldwork before and often alongside the collection of recordings. In most video-based workplace studies extensive fieldwork supports the analysis (Goodwin, 1995; Hindmarsh and Pilnick, 2002; Koschmann et al., 2007; LeBaron and Koschmann, 2003; Mondada, 2003; Sanchez Svensson et al., 2007, 2009). This fieldwork can be of a very particular nature. So, for example, to support the analysis of video materials gathered in emergency call centres, Jack Whalen trained as a dispatcher and learnt how to take calls in practice. In studies of control rooms and other 'technologically saturated' environments, extensive fieldwork is required to help make sense of the resources available to participants. For example, in a study of surveillance in the London Underground a principal resource for this work was the CCTV system. The CCTV system in each station control room provided access through a number (6–8) of monitors to most areas of the station, often from over 100 different cameras. Video data collection drew on a fairly standard framing of a shot including the monitors and the area where a principal member of staff responsible for managing the system – the station supervisor – was located (see Figure 14).



Figure 14. Video materials collected in station control rooms focused on the use of CCTV images. Understanding how staff made sense of the images required fieldwork concerned with the operation of the system, the contingencies of the setting and the kinds of problems faced by staff and how these were typically resolved.

This mid-shot framing provided access to most of the images and it was often possible to see features of the scenes that the staff were sensitive to. So, for example, the recordings support an analysis of how, after an incident is reported, the supervisor tracks and manages the incident through the use of the cameras and the monitors (Heath et al., 2002a; Luff et al., 2000). However, there are also cases that are more complex, the analysis of which required resources not immediately available in the recordings. Supervisors would often seem to discover events and problems just from a glance at a screen or as they switched from one view to another. Through fieldwork, it emerged that staff drew on a range of resources to make sense of the images and also to coordinate responses to the problems they discovered. Particular incidents occurred in specific locations, or at common times of the day, or could be identified from features not directly visible on the screens. Fieldwork provided a generic account of these. To understand how these were identified in practice and how staff coordinated a response with their colleagues, required gathering further materials, including details of the controls of the CCTV system, how it was used to handle typical cases (and also exceptional ones), problems that could occur with the technology, and how these were managed. Data were also needed about the other systems that are available in the control room, alarms that occur, communication resources and the various documents that are used in the setting, including logs, but also lists on the walls and information recorded on whiteboards.

Given the ubiquity of computer systems, it is often necessary to gain a detailed understanding of the operation of the technology, when and how it is used, and also the competencies that underpin their use. Some of this understanding can be gained through typical approaches to fieldwork, through informal interviews and collecting example documents, for example. But it often requires reading system manuals, obtaining details of the interface and recording sequences of activity displayed on a computer screen. It may also require some more formal training in the use of the technologies. Even after gathering such materials, it may still be necessary to draw on the participants' expertise to understand particularly complex fragments of activity. It may be that the video material collected can be useful in this regard. Fieldwork can be supplemented by reviewing

tentative analyses and selected fragments with the participants themselves; for example, to recover what might have happened in a particular case and whether what has been recorded seems typical or not.

Video can provide unprecedented access to a domain and can provide an invaluable resource that can be subjected to repeated scrutiny not only by the researcher but also by colleagues. A convention has emerged for video data collection that seems to support this kind of research activity, particularly the analysis of focused interaction. However, even in domains that are not that complex, gathering video recordings needs to be supplemented with fieldwork, not only observations to help make sense of the social organization of an activity but also the technical resources and competencies of the participants. Such fieldwork not only supports the analysis of audio-visual materials but can also inform how materials are gathered, suggesting suitable locations as to where to record and where to position a camera, whether this is for a conventional shot or to gather supplementary or alternative materials. Just as the conventional framing of a shot may need to be supplemented with additional materials, the analysis of video may need to draw on resources not immediately available in the recordings (see Schegloff, 1992).

Discussion

In a widely cited article, Barley and Kunda (2001) criticize recent social scientific studies of work, particularly within organization theory. They argue that conceptual development is hampered by a dearth of detailed studies of work. To move away from imagined or 'petrified' images of work they suggest some methodological requirements for studies to 'bring the work back in', in part, through a return to fieldwork and ethnographic studies. Naturalistic studies of everyday activities, supported by video recordings, would seem to be one way to address these concerns, not just for organizational studies but for other areas of the social sciences. Video offers unprecedented access to domains and activities and provides resources that can be repeatedly viewed that are open to review and analysis by others.

However, video presents a range of problems for analysts; for example, how it can be rendered in a suitable form to support analysis, how to select and transcribe particular fragments. It can also be challenging to present and analyse drawing on video materials, not only in traditional paper texts but also in talks and through electronic media. However, before these problems can be faced data need to be collected that are of suitable quality to be analysed. This, of course, relies on gaining access to a setting and obtaining the appropriate ethical approval, but it also requires capturing data that can be subjected to repeated scrutiny. To obtain data that are amenable to analysis the importance of what might seem mundane choices, such as at what height to place the camera, where to point the lens and how wide to set the focus, can be critical for the subsequent analysis that can take place. We have suggested that one way to address these problems is to adopt, at least at first, a conventional framing for the shot. A framing that is stable, provides some detail of the activities being considered, but also features of the surrounding environment, principally the key participants undertaking those activities. Such a framing can help with the subsequent selection of fragments to scrutinize, transcription of fragments and with the presentation of analyses.

Researchers who have drawn upon the stable mid-shot have developed a rich corpus of studies in a very wide range of domains, not just a great variety of workplaces, but also semi-public settings like markets and auctions and conduct in museums, galleries and science centres (Heath et al., 2002b; Hindmarsh et al., 2005; Vom Lehn et al., 2001). It can even serve as the principle resource for activities of more transient activities; for example, of driving, where a fixed camera angle has been utilised to capture the activities of the driver and front seat passengers in cars (Haddington, 2010; Haddington and Rauniomaa, 2011; Laurier and Brown, 2008; Laurier et al., 2008). These have revealed the ways in which activities are produced through social interaction with regard to features of the environment and included analyses of how the use of artefacts is shaped by and shapes the interactions between participants. In a wide range of settings this framing facilitates an analysis not only *that* a conduct is shaped or transforming with regard to others or features of the environment but *how*. Despite being rather limiting, the shot can open up for analysis how activities are accomplished from the participants' perspective and how trajectories of action emerge in a setting. However, by being fixed, this stable shot provides little access to activities of others not within its circumscribed focus or to details of features of a setting. Supplementing the conventional framing with other camera angles can open further details to scrutiny, but can introduce problems both in data collection and analysis.

Of course, there are many studies where the choice of the mid-shot framing, whether augmented or not, will be inappropriate. Even if the analysis aims to focus on the detailed production of activities, if those activities cannot be localized it may be difficult to gather recordings using this framing to capture the phenomenon of concern. Although in many activities there are locales where participants gather around stable objects, there are ones where the activity is accomplished while the participants are in motion (e.g. while walking) and the camera may need to follow them. Despite some technological innovations in this area with wearable cameras (e.g. Lahlou et al., 2009), whether these are adopted, or the more straightforward option of the roving camera, researchers are likely to have to analyse materials that are less settled.

Whether video materials are gathered using the conventional framing or not, apart from a few phenomena of concern, subsequent analysis will require some understanding of the domain. The collection of video recordings rarely is undertaken without some accompanying fieldwork; interviews and general observations of the domain being fruitful for selecting initial locations and activities to focus on, for example. The details that video recordings can make available can also require that fieldwork of a particular kind is required. This may in part account for why one of the seeming advantages of video – of its potential for reusability – rarely comes to pass. Given that the resource can be repeatedly reviewed, it would seem that video data collected for the purposes of one project could not be used in others, even where researchers adopt different analytic standpoints. However, even when the focus is on the production of fine details of localized activities, the need to understand these with regard to the circumstances facing the participants when they were recorded can undermine the potential for reuse. As in the cases where video is collected to illustrate a prior analysis or contribute to the development of some documentary record, the ways the data are collected are intimately tied to the analytic orientation that frames the research.

For the past 30 years, researchers who have sought to analyse the moving image have most commonly used video technology, and in this time there have been a number of developments of the technology. Not least are the number of formats that have emerged; from U-Matic, VHS and Betamax tape formats in the 1970s, through Video8 and Hi8 in the 1980s, and more recently DVC, MiniDV and digital formats for hard disc camcorders. Each has produced higher fidelity materials and typically has become simpler to handle, copy and manipulate. Despite presenting the possibility to open up more details of the activities being available to scrutinize, these enhanced technical capabilities do not resolve most of the problems faced by the social scientific researcher. Using multiple high-definition cameras and combining visual (and audio) data from different sources provide greater access to the setting, to the details of an activity, and can allow researchers to reveal the interrelationships between distributed activities. However, when analysing these, the researcher still has to consider the ways in which these multiple views resonate with the perspectives of the participants being recorded, how what is visible in the data is visible to them. It also remains a challenge to warrant the status of additional resources gathered to augment a study and how the analysis of these can be best integrated with the video recordings.

While video has helped in the development of a body of work that has begun to reveal the detailed ways in which many everyday activities are produced and made sense of through social interaction, there are still a great many unresolved analytic and methodological challenges facing researchers who utilise audio-visual materials. Where there is methodological guidance and debate regarding the use of video recordings, in the social sciences it tends to be focused on ethical concerns about gaining access, the ways in which recording an activity might transform the conduct being recorded and how the gathered material reflects biases of those undertaking the recording. While these are no doubt important, there are other considerations to be considered when gathering video data. If this kind of resource is to be better utilised in the social sciences, we need to open up to debate the methodological consequences of other choices and strategies used in the collection and analysis of audio-visual materials. Even the seemingly technical and practical choices made when video materials are collected, like the selection of what to record, might need to be unpacked, discussed and reconsidered.

Funding

This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

References

- Barley SR and Kunda G (2001) Bringing work back in. *Organization Science* 12(1): 76–95.
- Birdwhistell RL (1970) *Kinesics and Context: Essays on Body Motion Communication*. London: Allen Lane.
- Brand SB (1976) For God's sake, Margaret: conversation with Gregory Bateson and Margaret Mead. *CoEvolutionary Quarterly* 10: 32–44.
- Brown B and Laurier E (2005) Maps and journeys: an ethnomethodological investigation. *Cartographica* 4(3): 17–33.

- Condon WS and Ogston WD (1966) Sound film analysis of normal and pathological behavior patterns. *Journal of Nervous and Mental Disease* 143: 338–347.
- Gesell A (1925) *The Mental Growth of the Preschool Child: 'A Psychological Outline of Normal Development from Birth to Sixth Year, Including a System of Developmental Diagnosis*. New York: Macmillan.
- Gesell A, Halverson H, Thompson H, Ilg FL, Castner BM, Bates Ames L and Amatruda CS (1940) *The First Five Years of Life: A Guide to the Preschool Child*. New York and London: Harper and Brothers.
- Gilbreth FB (1908) *Concrete System*. New York: The Engineering News Publishing.
- Gilbreth FB (1911) *Motion Study: A Method for Increasing the Efficiency of the Workman*. New York: D. Van Nostrand.
- Gilbreth FB (1973 [1908]) *Field System*. Easton, PA: Hive Publishing.
- Gilbreth FB (1974 [1909]) *Bricklaying System*. Easton, PA: Hive Publishing.
- Goodwin C (1981) *Conversational Organisation: Interaction between Speakers and Hearers*. London: Academic Press.
- Goodwin C (1995) Seeing in depth. *Social Studies of Science* 25(2): 237–274.
- Goodwin C (2003) Pointing as a situated practice. In: Kita S (ed.) *Pointing: Where Language, Culture and Cognition Meet*. Mahwah, NJ: Lawrence Erlbaum, 217–241.
- Goodwin C (2004) A competent speaker who can't speak: the social life of aphasia. *Journal of Linguistic Anthropology* 14(2): 151–170.
- Goodwin C and Goodwin M (1987) Concurrent operations on talk: notes on the interactive organisation of assessments. *IPrA Papers in Pragmatics* 1(1): 1–54.
- Greatbatch D, Heath CC, Campion P and Luff P (1995a) How do desk-top computers affect the doctor-patient interaction? *Family Practice* 12(1): 32–36.
- Greatbatch D, Heath CC, Luff P and Campion P (1995b) Conversation analysis: human-computer interaction and the general practice consultation. In: Monk A and Gilbert GN (eds) *Perspectives on HCI: Diverse Approaches*. London: Academic Press, 199–222.
- Greatbatch D, Luff P, Heath CC and Campion P (1993) Interpersonal communication and human-computer interaction: an examination of the use of computers in medical consultations. *Interacting With Computers* 5(2): 193–216.
- Haddington P (2010) Turn-taking for turntaking: mobility, time and action in the sequential organization of junction-negotiations in cars. *Research on Language and Social Interaction* 43(4): 372–340.
- Haddington P and Rauniomaa M (2011) Technologies, multitasking and driving: attending to and preparing for a mobile phone conversation in the car. *Human Communication Research* 37(2): 223–254.
- Haddon AC (1993) Diary 10 March 1898–25 March 1899. In: Long C and Laughren P (eds) *Australia's first films: facts and fables*. *Cinema Papers* 96: 32–37.
- Haviland JB (1993) Anchoring, iconicity, and orientation in Guugu Yimidhirr pointing gestures. *Journal of Linguistic Anthropology* 3(1): 3–45.
- Heath CC and Luff P (2000) *Technology in Action*. Cambridge: Cambridge University Press.
- Heath CC and Luff P (1992) Collaboration and control: crisis management and multimedia technology in London Underground line control rooms. *CSCW Journal* 1(1–2): 69–94.
- Heath CC and Luff P (1996) Convergent activities: line control and passenger information on London Underground. In: Engeström Y and Middleton D (eds) *Cognition and Communication at Work*. Cambridge: Cambridge University Press, 96–129.
- Heath CC and Luff P (2007a) Gesture and institutional interaction: figuring bids in auctions of fine art and antiques. *Gesture* 7(2): 215–240.

- Heath CC and Luff P (2007b) Ordering competition: the interactional accomplishment of the sale of art and antiques at auction. *British Journal of Sociology* 58(1): 63–85.
- Heath CC and Nicholls GM (1997) Animating texts: selective readings of news stories. In: Resnick LB, Saljo R, Pontecorvo C and Burge B (eds) *Discourse, Tools and Reasoning: Situated Cognition & Technologically Supported Environments*. Berlin: Springer-Verlag.
- Heath C and Vom Lehn D (2004) Configuring reception: (dis-)regarding the 'spectator' in museums and galleries. *Theory, Culture & Society* 21(6): 43–65.
- Heath C and Vom Lehn D (2008) Construing interactivity: enhancing engagement and new technologies in science centres and museums. *Social Studies of Science* 38(1): 63–91.
- Heath C, Luff P and Sanchez Svensson M (2002a) Overseeing organisations: configuring the environment of action. *British Journal of Sociology* 53(2): 181–203.
- Heath C, Luff P, Vom Lehn D, Hindmarsh J and Cleverly J (2002b) Crafting participation: designing ecologies, configuring experience. *Visual Communication* 1(1): 9–34.
- Hindmarsh J and Pilnick A (2002) The tacit order of teamwork: collaboration and embodied conduct in anaesthesia. *The Sociological Quarterly* 43(2): 139–164.
- Hindmarsh J, Heath C, Vom Lehn D and Cleverly J (2005) Creating assemblies in public environments: social interaction, interactive exhibits and CSCW. *Computer Supported Cooperative Work: The Journal of Collaborative Computing* 14(1): 1–41.
- Hockings P (1995 [1974]) *Principles of Visual Anthropology* (2nd Edition). Berlin and New York: Mouton de Gruyter.
- Jirotko M, Luff P and Heath C (1993) Requirements for technology in complex environments: tasks and interaction in a city dealing room. *SIGOIS Bulletin* 14(2): 17–23.
- Kendon A (2004) *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Knoblauch H, Schnettler B, Raab J and Söffner H-G (eds) (2006) *Video-Analysis: Methodology and Methods Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt am Main: Lang-Verlag.
- Koschmann T, LeBaron C, Goodwin C, Zemel A and Dunnington G (2007) Formulating the 'Triangle of Doom'. *Gesture* 7(1): 97–118.
- Kress G, Jewitt C, Bourne J, Franks A, Hardcastle J, Jones K and Reid E (2005) *Urban Classrooms, Subject English: Multimodal Perspectives on Teaching and Learning*. London: RoutledgeFalmer.
- Lahlou S, Nosulenko V and Samoylenko E (2009) SUBCAM technology as an instrument in psychological science. *Experimental Psychology* 1: 72–80.
- Laurier E and Brown B (2008) Rotating maps and readers: praxiological aspects of alignment and orientation. *Transactions of the Institute of British Geographers* 33: 201–221.
- Laurier E and Philo C (2006) Natural problems of naturalistic video data. In: Knoblauch H, Raab J, Soeffner H-G and Schnettler B (eds) *Video-Analysis Methodology and Methods, Qualitative Audiovisual Data Analysis in Sociology*. Frankfurt: Peter Lang, 181–190.
- Laurier E, Lorimer H, Brown B, Jones O, Juhlin O, Noble A, Perry M, Pica D, Sormani P, Strebel I, Swan L, Taylor A, Watts L and Weilenmann A (2008) Driving and passengering: notes on the ordinary organization of ordinary car travel. *Mobilities* 3(1): 1–23.
- Laurier E, Whyte A and Buckner K (2001) An ethnography of a cafe. *Journal of Mundane Behaviour* 2(2): 195–232.
- LeBaron C and Koschmann T (2003) Gesture and the transparency of understanding. In: Glenn P, LeBaron C and Mandelbaum J (eds) *Studies in Language and Social Interaction in Honor of Robert Hopper*. Mahwah, NJ: Lawrence Erlbaum Associates, 119–132.
- Luff P and Heath CC (1993) System use and social organisation: observations on human computer interaction in an architectural practice. In: Button G (ed.) *Technology in Working Order*. London: Routledge, 184–210.

- Luff P and Heath CC (2000) The collaborative production of computer commands in command and control. *International Journal of Human-Computer Studies* 52: 669–699.
- Luff P, Heath CC and Jirotko M (2000) Surveying the scene: technologies for everyday awareness and monitoring in control rooms. *Interacting With Computers* 13: 193–228.
- Luff P, Heath CC and Pitsch K (2009) Indefinite precision: the use of artefacts-in-interaction in design work. In: Jewitt C (ed.) *Routledge Handbook of Multimodal Analysis*. London: Routledge, 213–222.
- Luff P, Pitsch K, Heath CC, Herdman P and Wood J (2008) Creating affinities between paper and digital resources. In: *Proceedings of Social Interaction and Mundane Technologies*. SimTech, Cambridge.
- MacBeth D (1999) Glances, trances and their relevance for a visual sociology. In: Jalbert PL (ed.) *Media Studies: Ethnomethodological Approaches*. Washington, DC: University Press of America, 135–170.
- Meisner R, Vom Lehn D, Heath C, Burch A, Gammon B and Reisman M (2007) Exhibiting performance: co-participation in science centres and museums. *International Journal of Science Education* 29(12): 1531–1555.
- Mondada L (2003) Working with video: how surgeons produce video records of their actions. *Visual Studies* 18: 58–73.
- Mondada L (2007) Operating together through videoconference: members' procedures for accomplishing a common space of action. In: Hester S and Francis D (eds) *Orders of Ordinary Action*. Aldershot: Ashgate, 51–67.
- Murphy K (2005) Collaborative imagining: the interactive use of gestures, talk, and graphic representation in architectural practice. *Semiotica* 156: 113–145.
- Pink S (2004) *Home Truths: Changing Gender in the Sensory Home*. Oxford: Berg.
- Pink S (2006) *Doing Visual Ethnography: Images, Media and Representation in Research* (2nd Edition). London: SAGE.
- Sanchez Svensson M, Heath CC and Luff P (2007) Instrumental action: the timely exchange of implements during surgical operations. In: Bannon L et al. (eds) *The Proceedings of the 10th European Conference on Computer Supported Cooperative Work*. London: Springer-Verlag, 41–61.
- Sanchez Svensson M, Luff P and Heath CC (2009) Embedding instruction in practice: contingency and collaboration during surgical training. *Sociology of Health and Illness* 31(6): 889–906.
- Schefflen AE (1973) *Communication Structure: Analysis of a Psychotherapy Transaction*. Bloomington: Indiana University Press.
- Schegloff EA (1992) In another context. In: Duranti A and Goodwin C (eds) *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, 193–227.
- Suchman L (1997) Centers of coordination: a case and some themes. In: Resnick LB, Säljö R, Pontecorvo C and Burge B (eds) *Discourse, Tools, and Reasoning: Essays on Situated Cognition*. Berlin: Springer-Verlag, 41–62.
- Tulbert E and Goodwin MH (2011) Choreographies of attention: multimodality as a routine family activity. In: Streeck J, Goodwin C and LeBaron C (eds) *Embodied Interaction: Language and Body in the Material World*. Cambridge: Cambridge University Press, 79–92.
- Vom Lehn D, Heath C and Hindmarsh J (2001) Exhibiting interaction: conduct and collaboration in museums and galleries. *Symbolic Interaction* 24(2): 189–216.
- Whalen J (1994) *Making standardization visible*. Department of Sociology, University of Oregon.
- Whalen J (1995a) Expert systems vs. systems for experts: computer-aided dispatch as a support system in real-world environments. In: Thomas P (ed.) *The Social and Interactional Dimensions of Human-Computer Interfaces*. Cambridge: Cambridge University Press, 161–183.

- Whalen J (1995b) A technology of order production: computer-aided dispatch in public safety communications. In: ten Have P and Psathas G (eds) *Situated Order: Studies in the Social Organisation of Talk and Embodied Activities*. Washington, DC: University Press of America, 187–230.
- Whalen J and Vinkhuyzen E (2000) Expert systems in (inter)action: diagnosing document machine problems over the telephone. In: Luff P, Hindmarsh J and Heath C (eds) *Workplace Studies: Recovering Work Practice and Informing System Design*. Cambridge: Cambridge University Press, 92–140.
- Whalen J, Whalen M and Henderson K (2002) Improvisational choreography in a teleservice work. *British Journal of Sociology* 53(2): 239–259.
- Whalen M, Whalen J, Moore R, Raymond G, Szymanski M and Vinkhuyzen E (2004) Studying workspaces as a natural observational discipline. In: LeVine P and Scollon R (eds) *Discourse and Technology: Multimodal Discourse Analysis*. Washington, DC: Georgetown University Press, 208–229.

Biographical notes

Christian Heath is Professor of Work and Organization and leads the Work, Interaction and Technology Research Centre at King's College London. He specializes in video-based studies of social interaction, drawing on ethnomethodology and conversation analysis, with a particular interest in the interplay of talk, bodily conduct and the use of tools and technologies. He is currently undertaking projects on auctions and markets, medical consultations and operating theatres, and museums and galleries. A number of these projects also involve the design and development of advanced technologies including, for example, media spaces, trust systems, and tools to interweave paper and digital documents.

Paul Luff is Professor of Organizations and Technology at the Department of Management, King's College London. His research involves the detailed analysis of work and interaction and drawing upon video recordings of everyday human conduct. He has undertaken studies in a diverse variety of settings including control rooms, news and broadcasting, healthcare, museums, galleries and science centres and within design, architecture and construction. Over the past few years, he has been particularly concerned with the use of apparently mundane objects and how these support complex work practices. He has drawn from this research to inform a number of projects concerned with the design of novel systems, including advanced media spaces and technologies that can interleave the use of paper and electronic materials.