Terhi Korkiakangas
Challenges in Archiving and Sharing Video Data:
Considering Moral, Pragmatic, and Substantial Arguments

Abstract
Social science researchers are facing new challenges in data archiving and sharing. The challenges encountered for video data are different from those encountered for other types of qualitative data. I will consider these challenges with respect to the moral, pragmatic, and substantial arguments with which funding bodies justify data archiving and sharing. Throughout the article, I will draw on a recent Economic and Social Research Council funded project, "Transient Teams in the Operating Theatre," in which our research team video recorded work activities in the operating theatre of a UK hospital, thereby dealing with highly sensitive footage. I will consider how video data, on most occasions, cannot be archived for re-use by the wider research community, but how new avenues could be developed so as to benefit from further research on such "unarchivable" datasets.

1. Introduction

Contemporary social science researchers in the UK have faced new challenges with regard to archiving and sharing of research data. Funding bodies have introduced data archiving and sharing as strategies to promote preservation, re-analysis, and secondary analysis of data. In the social sciences, these strategies are regulated by the UK Data Service, a body, which is managed by the UK Data Archive and largely supported by Jisc (formerly, Joint Information Systems Committee), the University of Essex, and the UK's largest organisation for funding economic and social scientific research, the Economic and Social Research Council (ESRC).

The justifications for data archiving and sharing hinge on different arguments, which have moral, pragmatic, and substantial underpinnings. In short, the funding bodies imply that a researcher has a moral duty to deposit data for archives when the research has been supported with public money. This way, the impact of research could be maximised for the benefit of the public. Facilitating access to shared data has also pragmatic value. That is, when data are obtainable directly from the archives, further research can be accelerated, reducing the time and effort it takes to collect new data. Thus, data sharing is seen as an economic strategy. The substantial argument for data archiving and sharing follows suit of the development of "big data." Namely, the accumulation of data can expand novel research and enable researchers to tackle innovative questions and to visualise patterns across diverse datasets. The realisation of novel associations rely on constant data influx and can have significant implications: for example, big data flow has predicted infectious disease outbreaks in Africa by tracking peoples' movement patterns via mobile phone usage, as well as more common flu outbreaks by analysing Google search terms faster than an analysis of any hospital records would allow (Shaw, 2014). Recently, the National Health Service in the UK has initiated a plan for sharing medical records of the British

population for research purposes. This plan joins the big data revolution, as sharing medical information can help building large, diverse, and longitudinal data archives for pattern recognition in risk factors and treatments.

Research data are considered particularly valuable when they involve so-called hard-to-get data. These might involve data generated in sensitive environments, for instance, in medical contexts. Data sharing also avoids burdening vulnerable and over-researched populations by making effective use of what is already available. Indeed, social and economic research uses a wide range of data: national and international survey data collections, international databanks, census data, and various qualitative data, such as interview and focus group data. Such data have been successfully archived and already used for further research. For example, Bishop (2005a), an academic researcher and a research archivist at the UK Data Archive, has conducted a secondary analysis of the archived interviews about food and eating habits of the residents of Great Britain born between 1870-1908, 1915-1935, and 1930-1955. The interview data were originally generated as part of other projects by other researchers, and Bishop's inquiry was distinct from the original works by focusing on historical patterns in the contemporary consumption of processed food. This research also enabled her, as an archivist, to reflect on the actual process of secondary analysis to "more effectively address concerns of prospective depositors" (Bishop, 2007, p. 1).

Many researchers have expressed misgivings about archiving and re-use of qualitative data, notably interview data, which are often characterised as sensitive. In the last few decades, visual methods have become popular means of data collection, and video-ethnographic methods have been increasingly used for observational research on professional practices. Such data can be sensitive in different ways. In this article, I will consider the arguments for data sharing and archiving and illustrate some of the challenges that arise in relation to video recorded data. I do not intend to provide a comprehensive review of the issues around data archiving, re-use, and secondary analysis (or the nuances of these terms either); these discussions can be found elsewhere (e.g., Bishop, 2005b, 2009; Corti, Day, & Backhouse, 2000; Hammersley, 1997; Irwing & Winterton, 2011; Moore, 2007; Parry & Mauthner, 2005). Rather, I will show that working with video data places particular constraints that can be distinct from the challenges that arise in relation to archiving and sharing interview data, even if an interview is being video recorded. For example, filming clinical work in hospital settings is very different from interviewing clinicians or conducting a survey on clinical practices. I will discuss these differences closely with examples from my own research, linking them to the moral, pragmatic, and substantial arguments for data archiving and sharing.

2. Moral Argument for Data Archiving and Sharing
In a revised Research Data Policy, the ESRC notes that their plan for data sharing rests on the principles stated by the Organisation for Economic Co-operation and Development (OECD, 2007): "publicly-funded research data are a public good, produced in the public interest" and "publicly-funded research data should be openly available to the maximum extent possible." In the same policy, the ESRC states,

[W]e expect grant holders to generate scientifically robust data ready for further re-use through positive encouragement of the exploitation of the results of research supported by us, as well as other organisations with full respect to intellectual property rights. (Economic and Social Research Council [ESRC], 2013, p. 2).

These expectations are intimately linked to an attempt to maximise impact of publicly funded research. As the ESRC notes, data are "the main asset of economic and social research" and "key to an informed public policy" (p. 1). Such an argument constructs research data as a public good and implies an obligation on the part of the researcher to deposit their data for use by the research community. However, archiving and sharing cannot be justified simply in terms of a researcher's moral obligation to the funders. Notably, the extent to which research data can be archived and shared is dependent on informed consent provided by individual participants, and the UK Data Service acknowledges this. As such, the researcher's obligation is not only to their funders and the public, but also to their research participants.

In response to this, the UK Data Service encourages expanding the consent to be sought from participants. Their current recommendation is that the informed consent form should no longer preclude data sharing by an outright promise to destroy data after the project; this used to be a standard practice. Instead, they recommend a statement about data sharing in the original consent form, which apparently provides a just opportunity for participants to opt in or opt out of data archiving and sharing. The UK Data Service asserts that consenting to video-based studies can follow the same protocols as other research: "Audio-visual recordings or photographs can be handled by the same kind of consent procedures as other research materials" (UK Data Archive, 2014a). It is true, for example, that the ethical consent procedure should always include an opportunity to ask questions as part of informed consent. Indeed, the UK Data Service encourages "open discussion" to give participants the right to decide whether to allow data "to be used more widely by the research community for future uses" (UK Data Archive, personal communication, January 15, 2014). However, asking for additional consent raises ethical issues that, for one, relate to how we obtain data from participants and for what purpose.

2.1. Consenting to Data Re-use: Consenting to What?

Informed consent is the fundamental ethical requirement in research. It has transformed from being a one-time event to an open-ended process that is "continually open to revision and questioning" (Economic and Social Research Council [ESRC], 2012, p. 30). Participants retain a right to withdraw their data at any stage, without having to provide a reason. However, this negotiation process is no longer possible when new researchers obtain data from the archives without having direct contact with the participants whose data they wish to use: checking, revision, and questioning of consent cannot be done. Indeed, Broom, Chesire, and Emmison (2009) question how new researchers will contact participants after personal identifiers have been removed as part of data archiving.

When researchers are encouraged to request additional consent for data archiving and re-use, a question arises: What are participants consenting to? Any future re-uses or novel research questions remain unknown at the time of establishing initial consent.

Thus, unseen ethical challenges might emerge when data are re-used by other researchers. For example, new inquiries and research questions might be against participants' personal values or beliefs or involve other problematic aspects, and participants might wish to decline from taking part in such studies. As Parry and Mauthner (2004) note, participants have a right to object to what they feel is "inappropriate or derogatory use or *re-use* of the data" (p. 148, emphasis in original). Yet, it can become problematic to withdraw data from the archives and projects where data are already in use.

Some (e.g., Corti & Backhouse, 2005) believe that if consent for re-use by other researchers has been obtained at the time of the initial consent, this can safeguard against such challenges. Yet, this argument does not fully capture the complexity of asking participants' consent even for the primary research, which comes first. It is against the backdrop of this moment when a researcher sits down with a participant, explains the project, and presents a consent form, that we can unravel some of the difficulties in discussing additional consent, particularly in video-based research.

Recently, I was a researcher in an ESRC funded project at Imperial College London in which our research team video-recorded teamwork in the operating theatres of a UK hospital. We had to gain consent from the patients and the operating theatre professionals undertaking a surgical operation. Hence, for us, the process was doubly challenging. My colleague, a research nurse, took charge of getting consent from the patients. As a nurse, she had first-hand experience in talking to patients, and her nurse-status seemed to put the patients at ease as they waited for their operations.

So how did consenting happen in practice? My colleague met each patient individually in the waiting room. Normally this happened half an hour or so before a patient was due to go under general anaesthesia. Sometimes the opportunity to meet patients was only a moment before anaesthesia, as the list of patients occasionally changed. This moment in the waiting room was the first contact my colleague had with patients, and therefore, this was the first time these patients heard about the research. Due to patient confidentiality, and the way in which operations are planned, we would not have been able to obtain patient information beforehand so as to contact them about participation.

The fact that every patient approached consented to filming was a sign of great cooperation. But it was also a sign of vulnerability. The moment my colleague met the patients with a clipboard and a consent form was often delicate: the one thing on these patients' mind was the operation and their own safety. Thus, inviting their operation to be filmed as part of research was, in many ways, a lot to ask. One patient out of 20 was initially concerned that the footage would be used in a television show, and they were given extra assurance that it would not be. The patients were also explained that as their bodies would be mostly covered, their identities would be concealed when the operation was in progress. They were not the focus of the research: rather, we were interested in how the surgical team communicated during their operation.

It was challenging enough to explain the research thoroughly so that a participant understood the aims and implications, and to manage to go through the lengthy and

detailed consent form, sometimes with additional time-pressure. Considering the reality of a situation in which the potential participants are feeling preoccupied and information can be hard to take in, it seems almost unethical to be making further requests that are not easily explained or comprehended. Asking additional consent for archiving and re-use would be akin to a salesman technique: having gained consent to one thing ("got a foot in the door"), the researcher then asks for more. If a researcher cannot fully know, predict, and explain the implications and future uses of data (Greely, 2007), then it raises the question whether additional consent obtained from participants is fully informed and thus ethically sound.

2.2. Responsibility for Research Participants is a Moral Obligation

The issue of trust between a researcher and a participant should not be underestimated. Mauthner (2012) reflects on her role as an interviewer when establishing a relationship with women whom she interviewed about a particularly sensitive topic, postnatal depression. Trust was a condition for these women to open up and, as Mauthner notes, to "speak the unspeakable, to tell me things that, as many said, 'I've never told anyone before'" (pp. 1-2). Trust is also central in video-based research, and I do not mean (simply) an interview that is being recorded, rather, I mean video recordings of people doing work activities while being filmed for research purposes.

Definitions of what counts as sensitive data can vary. Parry and Mauthner (2004) note that an assumed hierarchy in data sensitivity can emphasise some issues over others, for example, "someone's 'ordinary' life is less sensitive than, for example, being abused, having HIV, or having a criminal record" (p. 148). But filming someone's ordinary life or work activities can be sensitive to that person being filmed. I feel a strong sense of responsibility over the participants who took part in our operating theatre project. Like Mauthner, who was trusted with stories of the "unspeakable," we were trusted access to events that sometimes passed unnoticed by the surgeons and the nurses and to behaviours that were sometimes beyond their awareness. Video-based research is not just observing or listening; many observational studies are conducted without video recording and can be less intrusive. Rather, video-supported research produces a tangible record that can be replayed, slowed down, paused, and zoomed in so as to attend to the most detailed aspects visible in the record. Quite literally, a participant's behaviour can be put under a microscope.

Video-ethnographic research can have some similarities with documentary filmmaking. However, documentary films routinely expose people's lives and realities more openly than does a piece of academic research. In ethnographic research, the protection of participants is overriding and reflects a different ethical framework from the journalistic ethics which underpins documentary film production (Koehler, 2012). In documentary films, misportrayal and misinterpretation might be introduced in post-production deliberately, for the purpose of viewer entertainment. This can be harmful for those that have been filmed. Consider the recent high-profile UK documentary program, *Benefits Street*, which followed the residents of James Turner Street in Birmingham. Many of the residents were unemployed with their only source of income coming from social security benefits. Apparently, the program had been portrayed as a look into the community spirit on James Turner Street. Yet, in an

article to *The Guardian* one of the featured residents, Deirdre Kelly, revealed a particularly strong reaction to the broadcasted show:

[W]e couldn't believe what we were watching. We went mad. People growing drugs, smoking drugs, shoplifting. That is not what our street is about. Half the people they showed don't even live in our street. (Aitkenehad, 2014)

Could video data sharing, in the name of academic research, project similar risks of harm to participants? Probably not in these measures. Namely, access to archived audio-visual recordings and disclosive documents are managed by the UK Data Archive through a secure access control and granted only to genuine researchers, who will have to justify their request for some piece of data. Yet, researchers are in the position to use video approaches creatively, and visual data can offer new perspectives to previously reported findings or arguments (Rose, 2012). As such, a video-based researcher, like a documentary filmmaker, can use the "power of images to fascinate" and "to entertain" (Butchart, 2013, p. 684). The danger of inadvertent misinterpretation remains when video data are shared, and I will return to this issue.

3. Pragmatic Argument and Practical Challenges in Data Archiving and Sharing
The pragmatic argument stresses the economic value of data archiving and sharing. Recruiting participants for research projects can be difficult and data collection itself is often time consuming. Indeed, on the UK Data Archive website it is stated:

Collecting data from surveys, questionnaires or interviews for one study is a painstaking process. Providing that accurate records have been kept, data that have been collected for one study can be analysed again for an entirely different piece of research. (UK Data Archive, 2014b)

The practical benefit of data archiving is that a bona fide researcher can relatively quickly accelerate new research when having access to deposited data. This can maximise the usability of data generated through a "painstaking process" by someone else.

Collecting video data is no doubt painstaking. The permission to video record can be difficult to obtain, in the first place: filming feels invasive. While researchers can assure confidentiality in data handling and storage, which might elicit participants' trust in the researcher, video recording can make participants feel as though they are under surveillance. In healthcare contexts, filming can have particularly strong implications to how researchers, and ultimately, the general public, see the healthcare organisation caring for patients. In cases of litigation, data collected for research purposes cannot be withheld in the court of law, which overrides any grants of confidentiality that have been given to participants as part of informed consent.

Further, complete anonymity cannot be guaranteed in video-based research: people are often visually recognisable in videos, and the researcher will know the identities of those that have been filmed. However, anonymity can be protected by means of concealing the participants' identities, which is particularly important in dissemination activities, for example, in presentations and publications that use still images from videos. Yet, participants' visual recognisability creates practical

dilemmas for archiving and sharing video data with other researchers.

The archiving protocol encourages data anonymisation to maintain confidentiality in all archived data. However, anonymisation raises different dilemmas when considered in relation to the original data sources or their representations, such as transcripts. The UK Data Archive has provided tips for anonymisation, yet these only partially recognise the challenges when it comes to video-based research. When researchers archive interview transcripts, the use of pseudonyms and the alteration of other identifiable markers might suffice to protect participants' identities. However, it is rarely possible to transcribe the entire original video data source, namely because of the multimodal character of video that requires multimodal transcription (for more on multimodal transcription, see Bezemer, in press; Bezemer & Mavers, 2011). Should video data—the original information source—be visually anonymised by posterising images or by blocking out the eye-region of faces, this would seriously compromise the re-usability of data. The audio should also be anonymised to disguise participants' voices, and this further obscures the information source through voice alteration. Thereby, the analysis of social interactions might become extremely limited from audio-visually altered data. What is more, some contemporary editing software can revert visual alterations back to the original format, which poses a serious risk for the participants' identities to be revealed after data have been archived. The UK Data Service acknowledges that data altering techniques are also challenging to apply to large data files, and some researchers (Derry et al., 2010) have proposed that making samples of video data available is more feasible. Thus, the pragmatic dilemmas associated with video data archiving and sharing are to do with the laborious nature of data alteration and the technological difficulties in achieving it.

4. Substantial Argument for Data Archiving and Sharing

Data accumulation alone does not enhance either knowledge or practice. Researchers must ask suitable questions and use or design appropriate approaches for further analysis of the data deposited. The research questions and methodologies are intimately linked to the researcher's ontological and epistemological assumptions. Floridi (2012) notes an epistemological problem with the accumulation of data that relates to the recognition of small patterns, namely to be able to delineate "where the new patterns with real added value lie in their immense databases and how they can best be exploited for the creation of wealth and the advancement of knowledge" (p. 436).

As different research paradigms guide both primary and secondary analyses of data, it is important to consider their implications for archiving and sharing video recordings as "data." While the realist paradigm postulates objects in the world which are empirically observable and discoverable, the constructionist paradigm postulates constructed objects which are not directly observable, but become discoverable within a social and linguistic context. What kinds of research questions could be examined through video recordings; what are the pitfalls of asking new research questions from video data generated by other researchers?

4.1. Is Video an Objective Record?

A video record captures a limited view of filmed events. Notably, the camera lens is restricted to having "no peripheral vision, limited mobility" and a "narrow angle view" (Jewitt, 2012). The recorded events are always situated in wider contexts and video can only offer a partial representation of what is happening and why. In an example of a video-based research in exhibit halls, Dicks, Mason, Williams, and Coffey (2006) filmed children playing with "Kugel Ball," a heavy revolving granite sphere. As the authors were trying to understand the interaction of two girls playing with the Kugel Ball, they first thought the girls were pretending that the object was a planet. Only after interviewing them, the authors found out that the girls had been interacting with a "wishing ball." Thus, video cannot be held as a portrayal of some objective reality or claimed to have a "truth-telling function" (Butchart, 2013, p. 678). What we look at through the lens of our camera, or indeed what we manage to record, does not mean we have "got it right"—whatever that "it" is from which implications are drawn. This is an important consideration.

In our operating theatre project, we used wide-angle video cameras to capture a broad view of the operating theatre. That is, we filmed the happenings in the theatre as much as could be fit in the lens, without actively focusing on anyone or anything in particular. Thus, by zooming out we could even capture a glimpse of the adjacent preparation room from where nurses picked up instruments and supplies; we stepped back to observe without deciding, at the time of filming, what was important. Yet, the centre stage was nevertheless the operating table and the cameras were always pointing that way. Thus, a selection was made in the camera orientation and, consequently, many events were not captured, limiting the documentation of a broader context.

As the Kugel Ball example shows, what is not captured in the camera will remain unknown to the viewer, reader, or other audiences. Sometimes inferences made from videos can have substantial implications, for example, when new researchers code (as "snap shots") and rate participants' behaviours for levels of competence in their work. Consider the following example from our operating theatre footage, which illustrates how easily such inferences can be made. Figure 1 shows an anonymised still image of a scrub nurse, whom we call "Rose," standing by the instrument trolley. It is the scrub nurses' task to pass instruments to the surgeon throughout the operation and to guard and keep a track of where the instruments are. Rose is positioned on the left of two surgeons, who are conducting a keyhole operation and standing side by side by the operating table.
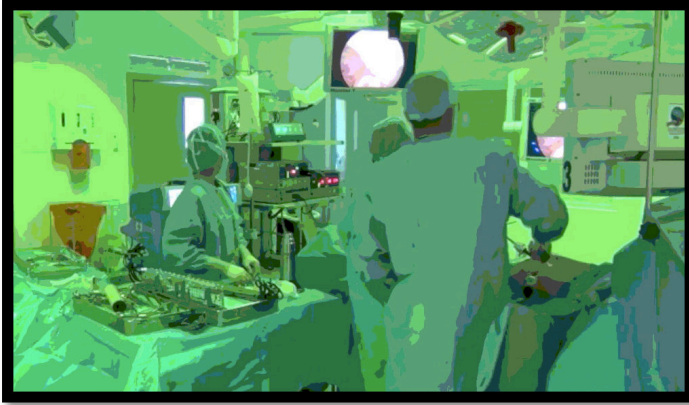
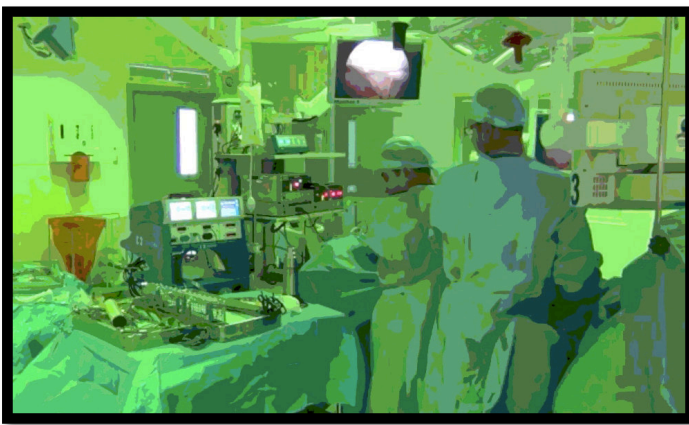Figure 1. Scrub nurse by the instrument trolley.

Figure 2. Scrub nurse away from the instrument trolley.

In Figure 2, the two surgeons turn to Rose, indicating their need for an instrument from the trolley. However, Rose is not there to assist: she has left the instrument trolley. It is unusual for a scrub nurse to be absent as it is the nurse's task to be available to assist with instrument exchanges. This moment became problematic. Both of the surgeons were holding laparoscopic instruments inside the patient's body and were compromised in their ability to move and reach to the trolley. It was also evident that the surgeons did not know where Rose had gone, as the equipment stack and monitors were obstructing a full view of the theatre.

As the consultant surgeon (on the left) called for "scissors," his tone marked annoyance at Rose's absence. Indeed, this particular episode had a ripple effect on the rest of the operation. The consultant displayed his unhappiness when Rose was delayed in her subsequent responses to his instrument requests, even though multitasking and liaising with other nurses requires that scrub nurses occasionally have to direct their attention elsewhere. If coded for technical competence and non-technical skills from this footage alone, Rose would probably score fairly low.

However, we (the original researchers) were present at the theatre and know more about the events leading up to Rose leaving her trolley. I was observing the operation by the wall opposite to the operating table, near the door (seen in the image) to the

preparation room. I saw and heard what was happening and was also caught in the middle of it. Earlier, the surgeon had requested a specific item, which was not available on the trolley. Rose relayed the request to a circulating nurse walking past the instrument trolley. The circulating nurse did not verbally acknowledge Rose's request (at least such acknowledgement was not audible), but continued walking out of the theatre. Although the circulator disappeared into the preparation room, Rose had been left waiting without a full confirmation that the circulator was going to retrieve the missing item, or that the circulator had even heard her. Several minutes passed, Rose was looking over her shoulder to monitor the circulator's return, and then finally left the trolley to look for her. Rose asked me to open the door into the preparation room; she was sterile and was not to touch anything beyond the instrument trolley. As I opened the door, Rose shouted into the room, calling on the circulator. It was right at this moment that the surgeon needed Rose and called out for "scissors."

In this situation, Rose leaving the instrument trolley could not be taken to indicate "inappropriate" behaviour on the job. This would be a serious case of misinterpretation. When we have played this clip to clinical audiences in workshops and conferences, their initial response has been a mixture of quiet laughter and headshaking: the scrub nurse must be "incompetent" for leaving the instrument trolley and the surgeons "in trouble." Rose was, in fact, problem solving in an acute situation, which required situation awareness and delegation for a missing surgical item to be retrieved. The primary researchers have knowledge and understanding that those who were not present in the situation do not have: cameras did not capture these events. Thus, video record is neither an objective record nor does it reveal the wider context in which events occur. This can distort a fuller understanding of why an observed situation has happened.

4.2. Complexity of Context

Many researchers (e.g., Dicks et al., 2006; Mauthner, 2012; Mauthner & Parry, 2010) have expressed criticism towards data archiving and re-use. These include: (a) risk of imposing ill-fitting research questions to data, (b) data can become decontextualised, (c) data can be misinterpreted, (d) risk for anonymity of participants, and (e) possible breach of trust between participant and researcher on data collected. In particular, the issues related to the context of data are problematic. Any assumed objectivity endangers data becoming decontextualised (Dicks et al., 2006; Mauthner, 2012; Mauthner & Parry, 2010) from the original context of generation, or treated as common currency (Hammersley, 1997), being treated as if data constitute an unproblematically transferable good between researchers. New researchers are inherently distanced from the data they did not obtain form the field but from the archives. In the case of video data, this can pose specific limitations for secondary analyses based on what is visually seen.

Indeed, many video researchers find it difficult to work with someone else's data; in video, the unfolding events are not apparently clear the way they might be to the original researcher. A good example of such situations are "data sessions" in which visiting social science researchers are watching clips of your video data with you. Often, a good deal of time is spent on explaining the context of the events; these are

not immediately available to others. We have shown several anonymised video clips in roundtable meetings and workshops involving academic colleagues and clinicians, and the context had to be discussed even if the data fragment lasted a few seconds or so.

In one example we showed during a workshop, a surgeon suddenly notices that a suction machine is not working and asks, "Is the suction working?" The apparent equipment failure necessitates that circulating nurses go and fix the problem; but also—at the more detailed level of interaction—a question (such as the one uttered by the surgeon) makes an answer to it relevant. However, the surgeon receives no immediate response to his question, which we found to be notable. Approximately 12 seconds later he turns to look over his shoulder, apparently in order to prompt a response from the nurses who had gathered around the suction machine to address the problem (see an anonymised still image in Figure 3). Another 11 seconds later, he turns to look again as he has received no verbal update about the situation.
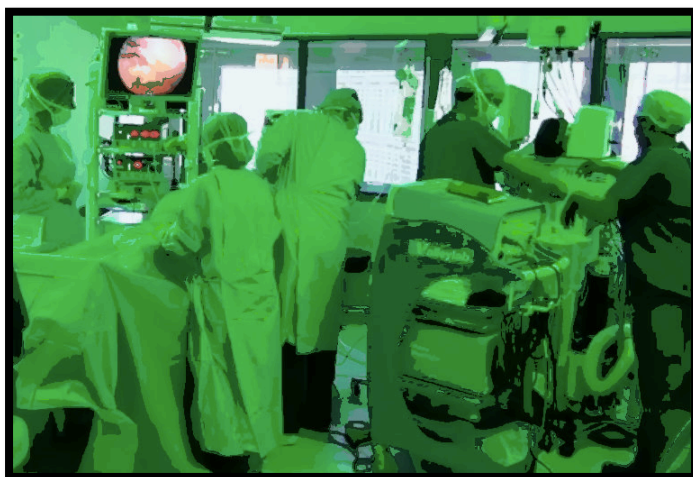

Figure 3. Surgeon turns to elicit a response from nurses.

As we played this clip in the workshop, we did not direct the audience's attention to anything in particular. First, we briefly explained the situation at hand (i.e., surgical operation is in progress and the surgeon experiences a failure with suction; then we identified the roles of the clinicians in the film) and simply asked them to notice anything that we could discuss in relation to communication. Somehow we, the original researchers, assumed that the lack of response from the nurses would be striking and obvious to others viewing the clip, but it was not—and that was interesting.

Instead, we received questions about the layout of the theatre, the tasks the participants were engaged in, did they know each other's names, and the like. In order to make sense of this short fragment, our audience yearned for much more information than what was shown to them. Even the seemingly "obvious" (obvious to us) fact that the suction machine was placed behind the surgeon's back and hindered his visual access to what was happening (and thus stressed his need of verbal updates), was not brought up by our audience. While videos do not reveal realities, also noticing is much more complex than it seems (Erickson, 2010). Visual—let alone analytical—observations do not always jump out of the data as givens.

Ethnographic data are always connected to a particular place, time, and wider societal ecology, which make the generated video unique to that setting. Yet, some research approaches make use of such information more than others, linking back to the underpinning research paradigms that have implications for the methodological choices. A researcher conducting content analysis, for example, might be interested in using rating scales as tools for coding and quantifying phenomena from videos, such as communication events during teamwork in the operating theatre. While these enable inferences to be made of the observed phenomena, they are limited in explaining relationships and why the phenomena might occur. Such coding does not necessarily need much contextual information, yet is limited in understanding what can be said about the coded behaviours beyond what is visually available.

The requirements for data archiving and sharing are accompanied by an encouragement to submit contextual information alongside metadata (i.e., searchable cataloguing information). But what information should the primary researcher include, at what level of detail, and how? Field notes are an essential part of video-ethnographic research that often describe and explain events and details which are out of camera view. The notes can include participant interviews, notes of informal chats in corridors, and other inputs that support or sharpen later analyses. Pink (2001) stresses the importance of noting down information conveyed through different senses, not just the visual. These are crucial to understanding observed events holistically: what something was like. In the operating theatre, this means documenting aspects of the environment—sounds, bleeps, smells, temperature—in order to understand what it is like also for the clinicians one is observing.

Clearly, not all such features are video or audio recordable, thus video footage offers a partial representation of the actual environment. While written notes can capture and describe some of these experiences, it is hardly possible to document everything about the context into a complete "package," ready to be handed over. Dicks and colleagues question:

[W]ould we want to provide just "the facts" alone (bracketing out for a moment the debates and controversies this idea conveys), i.e., just hand over the data records together with summaries of contextual information? This would be the "just facts" approach. Or do we want to keep it as complex as possible, so if you want to use the data you've got to read your way through and around them linking back all the time to the contingencies of data-generation and field relationships? This would be the "messy" approach. (Dicks et al., 2006, p. 35)

As a response to these challenges, Dicks' research group in Cardiff University has developed a web-based hypermedia platform for the dissemination and storage of data, contextual information and findings, in a multisensory way. The platform facilitates the preparation of datasets for archiving by linking different multimedia and hypertext resources: for example, a chronological sequence of photo stills or video footage can be linked to digitised field notes, which explain and expand on the moments captured, anchoring them to situational contexts and to other events not captured by cameras. This appeals as a useful way of demonstrating interconnections within the dataset, and visualising the researcher's analytic sense-making practices for others.

Yet, any contextual information generated by the primary researcher is also partial: it is not possible to observe and document everything. The information that has been documented is constructed through the senses and understanding of the researcher, or from the information provided by others, such as participants, from their perspectives. As Erickson (1986) reminds us, the underlying paradigms and presuppositions held by researchers also impact on their narrative descriptions of any observed events, even if these are observations of "ostensibly the 'same' behaviour performed by the 'same' individuals" (p. 120).

In order to decide how much and what kind of contextual information to provide, we need to know the research inquiries to be undertaken in future. But as these remain unknown, it leaves us in a loop of not knowing what information to provide. Hyperlinking and other means of dataset preparation are useful yet time consuming. Indeed, in a survey of data sharing practices by a sample of 1,329 scientists, Tenopir et al. (2011) found that the leading reason for withholding data from archives was insufficient time (54%) to prepare datasets for archiving, followed by lack of funding (40%) to do so.

While data archiving presents problems for researchers working with a range of data, including experimental data, observational data, and survey data, it is clear that for those working with qualitative data, difficulties involved in dataset preparation are even more laborious, costly, and complex. New researchers can surely bring novel observations and insights into video data, and perhaps see something that the primary researcher had overlooked. However, the wider understanding of such observations rests almost exclusively with the primary researcher, who had access to the complex contextual information and the social and technical insights that can be difficult to convey to others through written documents or even through multimedia platforms. As such, the primary researchers are a resource in the research process, in their own right.

5. Concluding Thoughts

While the issues around video data are complex, I would not position myself against data sharing per se. In fact, it is very much in the spirit of video-based social interaction research to host regular data sessions, often with colleagues and fellows from other institutions. In these sessions, video data, usually in their un-anonymised form (subject to consent), are reviewed, discussed, and interpreted together. Thus, video-based researchers already actively engage in data sharing during the course of their research. The data sessions generate new insights, observations, and possibly new research questions for the researchers to explore. However, the nature of this sharing is ephemeral and the data stay with the principal generator.

My view is, rather, that it might be impossible to (ever) archive video data for sharing and re-use, especially when considered sensitive. When consent has been given for archiving and re-use of sensitive video data, many questions still remain: What implications this might have in the long term? How "informed" is informed consent when future re-uses and re-users are unknown? How does the addition of detailed contextual information relate to the potentiality of participants and organisations becoming recognisable? How can we make video data anonymous so that they retain usability for further research?

As a possible solution, we could rethink research funding from two perspectives. Firstly, as found in the survey by Tenopir et al. (2011), researchers need more time and money to be able to deposit their data for archives. Indeed, funds can already be requested for this, yet the time allocated for data preparation can be extremely short. For example, the ESRC award holders must offer their data in the archives within 3 months of the end of their grant. It is imperative that more time and funding should be allocated for the preparation of datasets to meet the funders' requests for archiving. For some video data, the practice of hyperlinking might work well, but requires substantial resources. Alternatively, funders might consider a specific avenue for the primary researchers to form new research teams (involving new co-investigators and researchers not related to the original project) so as to fund new research on existing datasets that cannot be archived (possibly due to the sensitive nature of the data). Thus, even if data cannot be deposited in the archives, this would create another route for the data to be re-used.

References

Aitkenhead, D. (2014, March 7). Deirdre Kelly, AKA White Dee: 'I would never watch a show called Benefits Street'. *The Guardian*. Retrieved from http://www.theguardian.com/tv-and-radio/2014/mar/07/deirdre-kelly-white-dee-never-watch-benefits-street

Bezemer, J. (in press). How to transcribe multimodal interaction? In C. D. Maier & S. Norris (Eds.), *Texts, images and interaction: A reader in multimodality*. Berlin, Germany: Mouton de Gruyter.

Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: A social semiotic perspective. *International Journal of Social Research Methodology*, *14*(3), 191-207.

Bishop, L. (2005a, March). *'Oot o' the groun' and intae a pot': Convenience food and choice in the 20th century*. Paper presented at the meeting of the British Sociological Association, York, UK.

Bishop, L. (2005b). Protecting respondents and enabling data sharing: Reply to Parry and Mauthner. *Sociology*, *39*(2), 333-336.

Bishop, L. (2007). A reflexive account of reusing qualitative data: Beyond primary/secondary dualism. *Sociological Research Online* (Special section on reusing qualitative data), *12*(3), 1-14.

Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, *44*(3), 255-272.

Broom, A., Cheshire, L., & Emmison, M. (2009). Qualitative researchers' understandings of their practice and the implications for data archiving and sharing. *Sociology*, *43*(6), 1163-1180.

Butchart, G. C. (2013). Camera as sign: on the ethics of unconcealment in documentary film and video. *Social Semiotics*, *23*(5), 675-690.

Corti, L., & Backhouse, G. (2005). Acquiring qualitative data for secondary analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *6*(2), Article 36. Retrieved from http://www.qualitativeresearch.net/index.php/fqs/article/view/459

Corti, L., Day, A., & Backhouse, G. (2000). Confidentiality and informed consent: Issues for consideration in the preservation of and provision of access to qualitative data archives. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(3), Article 7. Availa Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1024/2207

Derry, S., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., . . . Sherin, B. L. (2010). Conducting video research in the learning sciences: Guidance on selection, analysis, technology, and ethics. *Journal of the Learning Sciences*, *19*(1), 3-53.

Dicks, B., Mason, B., Williams, M., & Coffey, A. (2006). Ethnography and data re-use: Issues of context and hypertext. *Methodological Innovations Online*, *1*(2), 33-46. Retrieved from http://www.esds.ac.uk/news/publications/MIODicksetal-pp33-46.pdf

Economic and Social Research Council. (2012). *ESRC framework for research ethics 2010*. Swindon, UK: Author. Retrieved from http://www.esrc.ac.uk/about-esrc/information/research-ethics.aspx

Economic and Social Research Council. (2013). *ESRC research data policy*. Swindon, UK: Author. Retrieved from http://www.esrc.ac.uk/about-esrc/information/data-policy.aspx

Erickson, F. (1986). Qualitative methods in research on teaching. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 119-161). New York, NY: Macmillan.

Erickson, F. (2010). On noticing teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 17-34). New York, NY: Routledge.

Floridi, L. (2012). Big data and their epistemological challenge. *Philosophy & Technology*, *25*(4), 435-437.

Greely, H. T. (2007). The uneasy ethical and legal underpinnings of large-scale genomic biobanks. *Annual Review of Genomics and Human Genetics*, *8*, 343-364.

Hammersley, M. (1997). Qualitative data archiving: some reflections on its prospects and problems. *Sociology*, *31*(1), 131-142.

Irwin, S., & Winterton, M. (2011). *Debates in qualitative secondary analysis: Critical reflections*. Timescapes Working Paper No. 4. Retrieved from http://www.timescapes.leeds.ac.uk/assets/files/WP4-March-2011.pdf

Jewitt, C. (2012). An introduction to using video for research. Mode Working Paper 3. National Centre for Research Methods, London, UK. Retrieved from http://eprints.ncrm.ac.uk/2259/1/MODE_Working_Paper_3_Video.pdf

Koehler, D. (2012). Documentary and ethnography: Exploring ethical fieldwork models. *The Elon Journal of Undergraduate Research in Communications*, *3*(1), 53-59. Retrieved from http://www.elon.edu/docs/e-web/academics/communications/research/vol3no1/06koehlerejspring12.pdf

Mauthner, N. (2012). Are research data a 'common' resource? *feminists@law*, *2*(2). Retrieved from https://journals.kent.ac.uk/kent/index.php/feministsatlaw/article/view/60

Mauthner, N., & Parry, O. (2010). Ethical issues in digital data archiving and sharing. *eResearch Ethics*. Retrieved from http://eresearch-ethics.org/position/ethical-issues-in-digital-data-archiving-and-sharing/

Moore, N. (2007). (Re)using qualitative data? *Sociological Research Online*, *12*(3). Retrieved from http://www.socresonline.org.uk/12/3/1.html

OECD. (2007). *Promoting access to public research data for scientific, economic and social development*. Paris, France: Author.

Parry, O., & Mauthner, N. S. (2004). Whose data are they anyway? Practical, legal and ethical issues in archiving qualitative research data. *Sociology*, *38*(1), 139-152.

Parry, O., & Mauthner, N. (2005). Back to basics: Who re-uses qualitative data and why? *Sociology*, *39*(2), 337-342.

Pink, S. (2001). *Doing visual ethnography: Images, media and representation in research*. London, UK: Sage.

Rose, G. (2012). *Visual methodologies: An introduction to researching with visual materials*. London, UK: Sage.

Shaw, J. (2014, March-April). Why 'big data' is a big deal. *Harvard Magazine*. Retrieved from http://harvardmagazine.com/2014/03/why-big-data-is-a-big-deal

Tenopir, C., Allard, S., Douglass, K., Aydinoglu, A. U., Wu, L., Read, E., . . . Frame, M. (2011). Data sharing by scientists: Practices and perceptions. *PLoS ONE*, *6*(6), Article e21101. Retrieved from http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0021101

UK Data Archive. (2014a). *Consent / Consent in audio-visual*. Retrieved from http://www.data-archive.ac.uk/create-manage/consent-ethics/consent?index=5

UK Data Archive. (2014b). *About the archive*. Retrieved from http://www.data-archive.ac.uk/about/archive