IJDC | *Peer-Reviewed Paper*

Privacy Concerns in Qualitative Video Data Reuse

Rebecca D. Frank
School of Information
University of Michigan

Allison R. B. Tyler
School of Information
University of Michigan

Anna Gault
Law Library,
Supreme Court of Ohio

Kara Suzuka
College of Education
University of Hawai'i at Mānoa

Elizabeth Yakel
School of Information
University of Michigan

Abstract

In this article, we examine how data producers' and reusers' privacy concerns shape their views about data sharing and reuse in the field of education, with an emphasis on video records of practice. We find that data producers and reusers were concerned about the risks that qualitative data, and video records of practice in particular, present to themselves, their colleagues, and the subjects represented in the data. Specifically, they emphasized risks relating to the privacy the subjects – teachers and students who appear in the videos. In response to these risks, data producers have engaged in a number of strategies to minimize risk and/or mitigate potential harm including: (1) education and training; (2) using informed consent to facilitate and/or restrict data sharing; and (3) limiting data capture/production. We discuss the implications that our findings have for digital repositories, and for efforts to facilitate the sharing and reuse of qualitative video data in education.

Received 06 June 2017 ~ *Revision Received* 1 June 2018 ~ *Accepted* 1 June 2018

Correspondence should be addressed to Rebecca D. Frank, School of Information, University of Michigan, 3429 North Quad, 105 S. State St., Ann Arbor, MI 48109. Email: frankrd@umich.edu

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

 doi:10.2218/ijdc.v13i1.492

Introduction

Despite recent efforts to increase data sharing and reuse, including funding agency mandates for researchers to deposit their data with repositories (e.g. Association of Research Libraries, n.d.; Holdren, 2013), these activities remain problematic in practice (e.g. Alter and Vardigan, 2015). This is particularly true for qualitative data in fields such as education, where K-12 teachers and students feature prominently in video records of practice (e.g. Cheville, 2016; Frank, Suzuka and Yakel, 2016; Gilmore, Adolph, Millman and Gordon, 2016). In this article, we examine how data producers' and reusers' privacy concerns shape their views about data sharing and reuse in the field of education, with an emphasis on video records of practice. We ask the following research questions:

- What concerns do data producers and reusers have regarding the privacy and confidentiality of subjects seen in video records of practice and how are these concerns represented in the risks or potential harms they perceive?
- How do perceptions of risk or potential harm influence attitudes about data sharing and reuse among researchers and teacher-educators in the field of education?

In education, a record of practice is data providing a "detailed documentation of teaching and learning" (Bass et al., 2002). Records of practice depict activities such as teachers leading classroom discussions and students engaging in group work. They include, but are not limited to, video and audio recordings from classrooms, still images of lesson activities, examples of student work, lesson plans, seating charts, and demographic information about students and teachers (e.g. test scores, socioeconomic information, etc.). They are a "window into other teachers' practice without having to be there in person" (Zhang, Lundeberg, Koehler and Eberhardt, 2011) and have long been used in educational research and teacher education (e.g. Burleigh and Peterson, 1967; Cohen, Burr, Goetz-Haver and Morales, 2003).

Background

Qualitative research is common in many disciplines, including education. Until recently, research about data sharing and reuse has focused predominantly on quantitative data; however, despite challenges, qualitative researchers are also interested in data sharing and reuse (Bishop, 2009). In this section, we discuss digital records of practice in education, qualitative data sharing and reuse, and the privacy and confidentiality issues that limit qualitative data sharing and reuse of these digital records of practice in the field of education.

Digital Records of Practice in Education

In the field of education, records of practice – and more recently digital video records of classroom activities – have been used in a variety of ways by researchers, teacher-educators, and teachers. For example, one of the primary ways that records of practice

have been used is in the development of professional vision among pre-service and in-service teachers (Blomberg, Stürmer and Seidel, 2011). As such, teacher-educators and professional developers have used video to help both pre-service and in-service teachers learn to "notice and interpret classroom interactions" (van Es and Sherin, 2002). Video is particularly useful for the development of professional vision and for highlighting skills and techniques in order to improve teaching practice (e.g. Penn-Edwards, 2012; Rook and McDonald, 2012). Teacher-educators and professional developers can use videos with teachers to review the same event multiple times, edit videos to highlight particular elements, and have teachers view classroom interactions without engaging in the work of teaching themselves (e.g. Penn-Edwards, 2012; van Es and Sherin, 2002). For studies of educational practices and methods, student learning and experience are significant components of validating outcomes. These are often depicted in the video records of practice, thus allowing researchers to study students' interactions with teachers and other students as well as through their work with materials, such as assignments and assessments. Classroom video also allows multiple users to view and re-view specific classroom activities, identify teaching methods, and interpret student and teacher behaviors without constant classroom disruption, for teacher training and development as well as research (Burleigh and Peterson, 1967; Masats and Dooly, 2011). Despite these benefits, researchers, teacher-educators, and teachers experience difficulties in sharing records of practice, and video in particular, due to technological, social, and ethical issues.

Data Sharing and Reuse

There are both benefits and challenges associated with data sharing and reuse. Manhas et al. (2015) summarize the benefits of secondary analysis as:

> '(a) increased diversity, novelty, and complexity of research opportunities thereby exhausting analysis potential; (b) cost savings through economies of scale to benefit the public, funders, researchers, and trainees; (c) lessened risk of not discovering key findings in the data; (d) promotion of intra- and inter-disciplinary research allowing multifaceted analysis; (e) maximization of research participants' contributions by fully utilizing the data; (f) lessened future research and respondent burdens; and (g) validation of previous work.'

Regarding qualitative data, Corti (2007) argues that the benefits of reusing qualitative data are access to rich, often unique data that "capture lived experiences of the social world and the meanings people give these experiences from their own perspectives." Usually qualitative studies cannot be replicated, nor can data be recollected even if the instrumentation is the same.

Data sharing and reuse can also be challenging. Data reusers often cite the lack of some metadata or supporting information, such as data producer name or affiliation (Berg and Goorman, 1999; van House, 2002) or date (Zimmerman, 2007) to help the data reuser judge the quality of the data or whether they are relevant to his or her research objectives. Others have specifically discussed the challenges of qualitative data sharing and reuse, including selection, analysis, technology, legal, institutional, and ethical concerns (Carlson and Anderson, 2007; Derry et al., 2010). The loss of research context is another potential risk that qualitative data reusers face. "One of the most persistent arguments around re-using qualitative data is about whether it is possible to

reuse data outside of the original context in which it was collected" (Moore, 2007). This lack of contextual information can limit reusers' ability and desire to reuse data for research (Yardley, Watts, Pearson and Richardson, 2013).

Managing sensitive data is a challenge faced by quantitative and qualitative researchers alike (Bishoff and Johnston, 2015). Regarding video records of practice in education, the risks of sharing and reuse include: (a) loss of anonymity; (b) reputational damage to teachers who are perceived to exemplify 'bad' teaching practices; (c) breaches of promises of subject confidentiality; (d) violation of teaching material copyrights; and (e) misinterpretations of the data (Arafeh and McLaughlin, 2002; Yardley et al., 2013). We take up these issues of privacy and confidentiality in greater depth in the following section.

In the field of education, there are a number of organizations that promote data sharing and reuse of digital records of practice. These include repositories managed by professional organizations (e.g. Accomplished Teaching, Learning, and Schools (ATLAS) from the National Board for Professional Teaching Standards), university-based video archives (e.g. Video Mosaic Collaborative at Rutgers, The State University of New Jersey), non-profit organizations (e.g. Achieve the Core from Student Achievement Partners), and for profit organizations (e.g. Teaching Channel).

Video records of practice include both large collections of raw video data and smaller, highly-curated collections of video clips. An example of a large collection is the Measures of Effective Teaching Longitudinal Database Extension (METX) collection at The Inter-university Consortium for Political and Social Research (ICPSR), a collection that requires special permission to access and use. An example of a smaller, highly-curated collection of video clips that are accompanied by additional contextual information and complementary data is the Third International Mathematics and Science Study (TIMSS) Video Study collection, which is open and available via a website.

In all cases, these repositories manage video collections that were either created for reuse from the collection's inception (e.g. Virtual Learning Community (VLC), Teaching Channel, METX) or reuse was later made possible through further collection development efforts (e.g. TIMSS, Video Mosaic, ATLAS). The repositories we studied did not include videos that were collected as part of a teacher preparation program or other context involving videos created by students as part of their coursework.

Privacy and Confidentiality

Data reusers' ability to access and work with data is hampered by concerns over privacy and confidentiality. In this paper we focus primarily on issues of privacy, however issues of confidentiality also arise as mechanisms that can be employed to protect subject privacy. We define privacy as the ability of research subjects represented in the data to remain anonymous. Still, we recognize that anonymity is not synonymous with privacy. Anonymity is a mechanism for protecting privacy, but one "can enjoy privacy while lacking anonymity, and retain anonymity while losing privacy" (Doyle and Veranas, 2014). We draw upon qualitative research literature which describes privacy for research subjects as the right to be left alone, to be free from surveillance, to make private communications, and to have bodily autonomy ('Privacy', 2008). Confidentiality refers to how widely data are shared and what access restrictions are in place. Unlike privacy, confidentiality is not concerned with data anonymization practices but with the mediation of access based on perceived risks of information disclosure (Lagoze, Block, Williams, Abowd and Vilhuber, 2013).

"Participant privacy is universally recognized as the primary concern when making data available to secondary researchers" (Manhas et al., 2015). Data producers are often reluctant to share their data for fear that reusers "might disclose data to the wrong people or simply disclose erroneous data that [have] not gone through appropriate quality checks" (Carlson and Anderson, 2007). Although data sharing policies favor openness, confidentiality and privacy are recognized as reasons for limiting access to data in order to protect subjects (e.g. Arzberger et al., 2004).

Privacy concerns exist for both quantitative and qualitative data. Varied examples of privacy concerns related to quantitative data include: 1) epidemiologists' using electronic medical records need to guard against the risk of revealing Health Insurance Portability and Accountability Act (HIPAA) information (El Emam and Moher, 2013), 2) archaeologists consider masking GIS location information due to concerns about looting (Frank, Kriesberg, Yakel and Faniel, 2015), and 3) marketers with access to large amounts of consumers' trace data ensure against data breaches leading to the identification of individuals (Custers and Uršič, 2016). Quantitative social scientists have approached privacy concerns in different ways, such as creating restrictions around highly sensitive data, limiting the amount of sharing, and anonymization through coding or other means (Narayanan and Shmatikov, 2008).

Privacy concerns are also endemic in qualitative data where personal information as well as one's image may be captured. In fact, the strength of qualitative research is the ability to capture a depth of experiences and personal details in the data. Several mechanisms have been established to address privacy concerns with qualitative data including: 1) institutional, educational, and legal structures largely through responsible conduct of research training and informed consent regimes aligned with Institutional Review Boards or Ethics Panels, 2) repository restrictions and establishment of confidentiality rules, and anonymization. Arafeh and McLaughlin (2002) argue that these techniques are more expensive and time-consuming for qualitative data. Thus, the sharing and reuse of qualitative data present unique challenges with regard to privacy.

Research ethics

Respect for research subjects through informed consent is major way that data sharing and reuse is handled for qualitative data. Through participant informed consent, research participants know the study purpose and authorize the data to be used for primary analysis in its original context. Yet, informed consent can present problems for data reuse when it does not explicitly include permission for preservation, sharing, and/or reuse (Corti, Day and Backhouse, 2000). To mitigate these concerns, "consent for the long-term storage of and access to data use should be obtained in all fieldwork circumstances, where this is possible" (Corti, Day and Backhouse, 2000). Still, few qualitative studies provide for secondary analysis and data reuse.

An area of particular concern for video records of practice is the presence of children in classroom videos (e.g. Flewitt, 2006). There are difficult issues surrounding permissions and consents involving children, including who should provide consent and when. While questions about the scope of informed consent for data sharing and reuse exist for all subjects (e.g. Bishop, 2009; Yardley et al., 2013), there are additional concerns in the field of education regarding the ability of minors to give fully informed consent (or the ethical right of adults to give full consent on their behalf) (Andersson and Sørvik, 2013). Although it is the children who appear in video records of practice, it is their parents or legal guardians who must consent (Hammersley, 2015; Manhas et al., 2015). This consent may last in perpetuity, even after the student has reached the age of consent. Many researchers whose work might affect children argue that children must

be included in that research and consent process, while others question whether research for the public good overrides a child's right to privacy (Darian-Smith and Henningham, 2012; Joyce, 2011). There is also "growing tension between researchers who seek to enable child participants to speak in their own voice and regulators who seek to restrict some studies in order to protect children and their privacy" (Darian-Smith and Henningham, 2012). Because archived videos will be available long-term, questions arise about what provisions, if any, should be made to allow children to reevaluate access and assert their rights over the videos when they reach the age of consent (Moser, Chen and Schoenebeck, 2017).

In addition to the long-term effects of consent and data reuse on the children in the videos, the parents themselves may have broader concerns for themselves or their family as a whole. For example, in medical research, parents may also be concerned about physical or mental diagnoses harming the reputation of the family or increasing insurance costs if known (Manhas et al., 2015). In video records of practice, a parent's refusal to let their child be video-recorded for research as well as further reuse of the recorded data often involves changing the production parameters or post-production processing.

Likewise, professional teachers whose images and teaching are recorded in video records of practice and supplemental data have concerns about when, where, and how these data will be used, even when they have given their informed consent to be recorded (e.g. Arafeh and McLaughlin, 2002; Bishop, 2009). In addition to concerns about whether their teaching practice will be used as an exemplar of good teaching or as an example of ineffective teaching, they also worry about how the context of the data may be lost (Andersson and Sørvik, 2013; Corti, 2000). Similarly, Bishop (2009) found that research participants have concerns about "how much say they have in data interpretations and research conclusions." This is a reflection of the desire for both Wallace's (2008) namelessness and freedom from potential consequences, and for control over the access to data that may increase their risk of reputational harm.

Repository restrictions

Confidentiality considerations are often addressed with access restrictions. Determining which data to make available and whether to apply access restrictions to those data can be a difficult decision for repository staff without input from the data producers. Corti, Day and Backhouse (2000) describe several strategies for making decisions about access restrictions, including measures controlled by the repository and measures controlled by the data depositor. They note that "we must place a significant portion of the initial responsibility for allowing people to have access to data with the investigators." Eschenfelder and Johnson (2011) found a high degree of variation in access restriction needs across disciplines and data types, and suggested that policies are needed which "allow researchers to retain more control over deposited data" in order to encourage data sharing.

For example, ICPSR, a social science data repository, describes an approach to access restrictions that relies on a combination of risk review by trained staff, and input from data depositors (e.g. information about IRB determinations, informed consent, etc.) (ICPSR, n.d.). In some cases, the ICPSR guidelines note that anonymizing data "significantly compromise[s] the research potential of the data" (ICPSR, n.d.). As a result, the repository offers five levels of access for data that cannot be anonymized, ranging from secure online access to an on-site physical data enclave.

Anonymization

Anonymity should be considered along several dimensions that are kept unknown to the audience (e.g., names, linkable and nonlinkable pseudonyms, locatability, etc.) (Marx, 1999). As previously noted, researchers require context to reuse data. In the case of qualitative data, contextual information includes elements such as age, geographic location, and socioeconomic status, which can be used to identify research subjects when aggregated. Anonymization practices employed to maintain the subjects' privacy can obscure or erase these types of contextual information. Reusers argue that without context, the results of secondary data analysis are not trustworthy (Yardley et al., 2013). As such, there is concern among data reusers that "qualitative data cannot be used sensibly without the accumulated background knowledge which the original investigator acquired" (Corti, 2000).

This need for contextual clues can hinder anonymization, as video records introduce the possibility of inappropriate disclosure of subjects' identities and locations (Carlson and Anderson, 2015; Manhas et al., 2015). Anonymization measures for video records of practice include masking faces and muting speech of non-consenting participants in video, limiting demographic data, or censoring supplemental materials to obscure the identities of subjects (Mondada, 2014). However, these anonymization measures may hinder data utility for secondary analysis because it "is very difficult to hide the identities of the people and places that appear on a videotape and, in many cases, such measures alter the data" (Arafeh and McLaughlin, 2002). As a result, qualitative data reusers' needs may also not be met by anonymization because data will then lack the information that they need, and minimize the utility of the video data for reuse (Andersson and Sørvik, 2007; Moore, 2007). For example, in our research, privacy concerns arise in the video records of practice that depict teachers and K-12 students. However, the research and teaching goals that many of our participants describe cannot be accomplished with anonymized data.

Furthermore, anonymization of video is difficult. The tools supporting anonymity in video data may not obscure all aspects of the individual, leaving identifiable characteristics, such as body shape, motion, voice, and personal traits or actions, which introduce the potential for reidentification and trigger the concerns discussed above (Wallace, 2008). Data producer and reuser concerns with the reuse of video data are often related to the ability to make these connections between traits/actions and personal identities, which may prevent them from maintaining the anonymity of subjects represented in video data.

Methods

This study is part of a larger research project, the Institute of Museum and Library Services funded Qualitative Data Reuse: Records of Practice in Educational Research and Teacher Development project. In this article, we focus on 44 interviews conducted with researchers and teacher-educators who have reused digital video as qualitative data for research and/or teaching.

Participant Recruitment

Our sample consists of data reusers in the field of education, with an emphasis on mathematics education research and teacher education. We define data reusers as

individuals who use records of practice that they themselves did not produce for new research and/or teaching purposes other than the original intent of the data producer (Faniel and Jacobsen, 2010; Zimmerman, 2008). We consulted the research literature and attended conferences to recruit participants using convenience sampling to identify researchers, pre-service teacher-educators, and facilitators of professional development for in-service teachers. We also asked interviewees to nominate additional interviewees using a snowball sampling technique. Of those 44 interviewees, the majority identified either research or pre-service teacher education as their primary areas of reuse, and described themselves as university faculty members. The majority of our interviewees were also based at institutions in the United States.

Table 1. Primary areas of data reuse (n = 44 interviewees)

Primary Data Reuse	Frequency	Percentage
Research	17	38.64
Pre-Service Teacher Education	21	47.73
In-Service Professional Development	4	9.09
Personal Study	1	2.27
Data Producer With No Reuse Experience	1	2.27

Table 2. Primary interviewee roles (n = 44 interviewees)

Role	Frequency	Percentage
University Faculty	31	70.45
Graduate Student	5	11.36
Postdoctoral Researcher	2	4.54
Other Education Professional	6	13.64

During the course of the interviews, we asked interviewees about their data reuse practices, including their attitudes about responsible conduct of research, appropriate use of digital records of practice, confidentiality and ethics regarding data reuse, and attitudes about the role of repositories in the field of education (Yakel, Suzuka and Frank, 2018). These semi-structured interviews lasted approximately one hour. They were audio recorded and transcribed for analysis. This study was reviewed and deemed exempt by the Institutional Review Board at the primary author's university.

Analysis

We analyzed the interview transcripts using NVivo, a qualitative data analysis software. We developed an initial code set based on themes from the literature. We then added additional codes based on themes that arose during the initial coding and analysis process (Miles and Huberman, 1994; Saldaña, 2015). This approach to qualitative data analysis combined deductive and inductive approaches, and incorporated descriptive, analytic, and thematic coding, thereby allowing us to compare our data with existing themes in the literature and identify new themes arising from this study. Through this process of axial coding, we developed a final code set which addressed topics such as data reuse practices, data sharing practices, ethical and legal issues regarding data reuse, and participant attitudes toward data repositories (Yakel et al., 2018).

We coded the interview transcripts in two groups: interviewees whose reuse focused primarily on research, and those whose reuse focused primarily on teacher education, including both pre-service and in-service contexts. For each group of interviews, two coders worked independently coding the same transcript. We repeated this process until we reached an acceptable level of interrater reliability for independent coding. Using Scott's Pi, a statistic measuring interrater reliability for coding textual data (Holsti, 1969), we achieved a score of 0.712 for the researchers, and 0.732 for the teacher-educators.

After this initial round of analysis, we further examined the data, isolating several codes (e.g. ethics and legal, appropriate use, confidentiality, and responsible conduct) for additional analysis, which embraced an axial coding approach to identify relationships between and dimensions of the existing codes, particularly on the topics of privacy, confidentiality, and risk/harm. A single team member, with substantial input from the entire team, conducted the second level analysis of the coding.

Findings

We organize our findings according to the three main themes based on our analysis of the interviews: privacy, potential harms, and mitigation techniques. First, we examine privacy in qualitative data production and sharing, and investigate how our interviewees characterized those issues for the field of education. Second, we identify and discuss the main categories of potential harms that participants associated with sharing and reusing video records of practice. Third, we discuss methods that we found our participants have used to mitigate those harms, focusing on education and training, informed consent, and limitations on sharing.

Data Production and Sharing

We found that although video records of practice comprise a significant portion of qualitative data in the field of education, researchers and teacher-educators have experienced challenges in maintaining the privacy of subjects represented in the data when producing, sharing, and reusing video records of practice.

A majority (82%) of the data reusers in this study were also data producers. Our interviewees most commonly produced video records of practice that included teachers and students in the classroom, individual instruction, or interviews with teachers and students:

> 'I've collected videos from ten teachers' classrooms. I've been doing point-of-view observation. That's when teachers wear a camera while they're in the midst of teaching, and actually save short video clips after they happened using a remote. Then I interviewed them about each of the clips afterwards. In addition to that, I set up some stationary cameras and audio recorders. I've both done that using those interviews as an access to how teachers are thinking in the midst of while they're teaching' (Interviewee 029).

The researchers and teacher-educators interviewed also produced many other types of qualitative data, including text and still images from classroom activities, and materials that teachers produce when preparing for lessons:

> 'Primarily things that help us make sense of the video, like our copies of what the students are doing at the time, teachers' lesson plans, so we have a context for what's going on, pictures of things that happened in the classroom. So things like writing on the board, we'll snap a picture as well as have the video' (Interviewee 021).

This experience as both data producer and data reuser made many of our interviewees particularly sensitive to the privacy concerns in their work. While collecting these records of practice, researchers and teacher-educators balanced their needs against the privacy considerations of the teachers and students in the videos because our analysis indicates that video records of practice are uniquely valuable:

> 'You can't completely capture a phenomenon with text. That will give you a lot of sense of what it's about and you have to be very skilled to do that well, but you also want to be able to get live images of what it is that you're actually studying and trying to understand, and studying instruction is really an example of that' (Interviewee 001).

Because of the unique nature of the field of education, data producers reported persistent issues with their ability to protect the privacy of the teachers and students who appear in the records of practice. Unlike other types of research in which researchers can assemble groups of volunteers, or dismiss those who do not wish to be included in the data, our participants explained that since they are filming actual classroom lessons, they cannot dismiss students who do not want to be studied. Interviewees expressed an obligation to students, to provide them with equal learning opportunities whether they consented to be included as study participants or not:

> 'When they were not doing research, they just had to do their little legal documents and then send the kids away who didn't agree. I can't do that. We can't send kids out. That's not what we do as researchers. Children can't be punished for not participating. We're going to have kids in videos that are great that we can't show and we're going to have to figure out can we edit it so that we exclude them? Can we blur and alter voices? What can we do?' (Interviewee 32).

Rather than limiting initial data collection based on shooting around the students who opt out, data producers such as Interviewee 32 capture the classroom as completely as possible and edit those individuals out post-production. This ensures that they are able to capture as much usable data as possible, but also means that they may have to cut out significant events from their final data depending on who is captured in any given frame or clip.

Data producers also cope with the need to edit footage by triangulating among multiple recordings and adding in additional information, such as lesson plans or examples of student work. This result is the creation of complex data objects consisting of video, audio, still images, subtitles and/or transcripts. Triangulation among these different elements enables users to produce usable versions of significant classroom

events, while protecting the privacy of teachers and students who have opted out of the research:

> 'We've then edited the video. And yeah, that usually involves transcribing, sort of cleaning the transcripts, picking the camera and audio feeds that are relevant, having that sort of sliced together, subtitled and students blurred' (Interviewee 22).

Researchers and teacher-educators often rely on complementary data, such as lesson plans, classroom handouts, and copies of student work to provide additional context and make sense of audio and video records of practice that they produce:

> 'We are also using any handouts that are used ... And we also have field notes that were taken at the same time that the video was taken, that we sometimes use for creating data summaries, logs about what's happening in class time' (Interviewee 14).

This triangulation of these different types of data introduces additional privacy concerns. For example, triangulation increases the ability to identify individuals through data aggregation. Even when video and audio are edited to remove identifying features, lesson plans, student work, and seating charts can provide identifiable information.

Data sharing is another activity where privacy considerations arise. Video records of practice carry unique challenges because there are often close relationships between data producers, reusers, and subjects represented in that data. Our analysis indicates that this is particularly problematic when using video records of practice from teacher-education contexts taught by university faculty colleagues. For example, Interviewee 5 explained the unique potential for harm to subjects represented in her research because of close relationships and membership in the same professional community:

> 'The person/people in the video are going to be identifiable, right? And so, when I go off to a conference and I show a video of our classroom teacher, for the most part, at a national conference, the chances of the person would be recognized are pretty slim, right? I feel like it's a little bit different when you're a teacher-educator and the people doing the work are your colleagues and will be showing it in the venues in which you are going to be participating, right?' (Interviewee 5).

In cases where interviewees talked about sharing their data, concerns about privacy and anonymization still surfaced as issues that influenced decisions about how to present qualitative data for potential reuse. For example, Interviewee 40 described the steps that he took to anonymize his data:

> '[F]or us we're just very fortunate that we kind of got in before people realized things about privacy. But, at the same time I go through the painstaking detail of every piece of student work I put up there, I make sure to redact the last name so you can only see the first name and the videos, go through and make sure there's no clips that would embarrass students, clips that would reveal who the students are and so forth. It's really kind of– spent a lot of painstaking time going through and making sure that students can't be identified. You show the video in class and it sounds like, "Is that so and

> so, I know them from blah, blah, blah." It's always that kind of issue but we try to spend a lot time to make sure to keep the student's identity private even though they have signed off on it' (Interviewee 40).

In response to the challenges that arise in data production and sharing, our analysis revealed that data producers who shared their data and data reusers must consider the risks that result from lack of privacy in video records of practice, which may potentially harm subjects and users.

Privacy, Risk, and Potential Harm

Our interviewees articulated several types of potential harm connected to the difficulty of protecting the privacy of subjects in the video. These included: (1) personal harm resulting from culpability for data misuse by data producers and/or reusers; (2) reputational or economic harm to subjects represented in the data; and (3) harm through identity exposure for subjects who appear in video records of practice, potentially amplified by the close ties that often exist between data users and subjects. Since techniques, such as anonymization, decrease the utility of the video records of practice, these concerns were heightened.

Interviewees were concerned that violating the privacy of teachers and/or students would negatively impact their own careers, whether the breaches of privacy or confidentiality were intentional or not. They thought this applied to their own conduct as well as that of others to whom they gave permission to use the data. For example, Interviewee 5 explained that the misuse of data she produced would damage her reputation and could hurt her job security by preventing her from getting tenure:

> 'And so I think there's a bit of a challenge there I think, just in terms of thinking about, what would you want to be shown in a public venue. And particularly thinking about some of my own positioning as someone who's not a tenured track faculty member but eventually does want a tenure track position. So just thinking about that. I think there is a little bit of vulnerability about sharing things, which I think could be managed by how the ways in which you set up what can be done with the data' (Interviewee 5).

The vulnerability that Interviewee 5 experienced extends beyond responsibility for what she did directly with the data to include what others might do in the future with data that she had produced and shared.

The second type of harm we identified was harm to the subjects represented in the data. Our interviewees expressed concern that the video records of practice would be used against the teachers in the videos, causing harm to their reputations. They noted adverse consequences ranging from viewers making snap judgments about the skill level of the teachers, to state or district officials and school administrators using the records of practice to evaluate the teachers, set policy, or make funding decisions. For example, Interviewee 9 stated that although producing and sharing video records of practice could have tremendous benefits for the field of education, doing so ethically requires finding a way to assure subjects that their participation will not be used against them in the future, "If there really were a way for people to be assured that it was not going to be held against them evaluation-wise and all of that, that it would be so powerful" (Interviewee 9).

Producers and reusers also expressed a related concern that a video showing a teacher or student on anything less than their best day would be taken out of context and used as a negative example, publicly shame them, or to evaluate their performance, thereby endangering their employment or the student's status. This framing highlights the commonly held attitude that teachers and students are particularly vulnerable, and in need of protection, "I think we all tend to be pretty protective of the teachers who are in our records of practice" (Interviewee 21).

Recognizing the vulnerability of subjects represented in video records of practice was an important theme among our interviewees. Interviewee 2, for example, explained that with video records of practice in education, breaches of privacy tended to affect groups who already lacked power or agency, such as children and teachers, "[a]nd when things go wrong, it's usually populations who have been stripped of voice and I think that that risk factor is real and needs to be accounted for" (Interviewee 2).

Similarly, our interviewees talked about how the risks to data producers, reusers, and subjects were heightened because they belonged to the same professional community. The audience for their work sometimes included the subjects who appeared in the records of practice. Interviewee 24 described an experience in which one of the subjects was in the audience during a presentation in which the audience member was featured in a video clip. As a result, she became more cautious about sharing video even in a classroom setting:

> 'The teacher was at one point in a conference when I was sharing some results of people looking at his video. So I am sensitive to the fact, like these are practicing teachers who have made their practice public and we can't ... I don't want my candidates to be over ... Or any teachers for that matter, being overly critical of what they're seeing, which is what teachers tend to do' (Interviewee 24).

In light of the three types of potential harm that we identified among our interviewees with regard to data production and reuse (i.e. harm to data producers and reusers, harm to subjects represented in the data, and lack of anonymity for subjects who appear in video records of practice), we next asked what methods interviewees had developed to protect themselves, their colleagues, and their subjects. We found that data producers and reusers mitigate harm through several approaches, which we describe below.

Mitigating Harm

In response to these risks, our interviewees employed a variety of techniques to mitigate potential harms to data producers, data reusers, and subjects represented in the data. Several interviewees emphasized that data in the field of education often depicts people, and that one of the primary responsibilities of data producers and reusers is the protection of human research subjects and the mitigation of harm:

> 'Well I think you have to remember these are oftentimes with humans because it's educational research, so you know anything that could potentially come back around and hurt a human you have to really protect against that' (Interviewee 16).

We organize our discussion of the techniques that our interviewees have employed to protect themselves, their colleagues, and their research subjects from potential harm resulting from breaches of privacy or confidentiality into three categories: (1) education and training; (2) using informed consent to facilitate and/or restrict data sharing; and (3) limiting data capture/production.

Education and training

One of the ways we found that our interviewees addressed their own potential culpability for misuse of data was through education and training about how to be responsible data reusers and producers.

Interviewee 44 described the steps that she took when reusing video records of practice in her teaching:

> 'The students, they have to be on campus to watch the videos so we can track that. If there are any handouts or things that I have the print out for them and they are numbered so I get them back. They all have to sign on our IRB the ethical use of materials saying that we have this embedded, and we also have that for each class or the way in which materials are allowed to be used and that's a court violation if they mess that up' (Interviewee 44).

She went on to explain that she talks with her students about the ways that video should be reused:

> 'I'm going to start with the idea of who owns the video and who decides what is the range of parameters for how their videos can be used. If there's specific and clear guidelines for that, we are very careful to follow those and we make sure the students know' (Interviewee 44).

Interviewee 24 explained that although she was not concerned about privacy issues with publicly available data, she still took steps to teach her students how to be respectful of the subjects represented in the publicly available videos:

> 'I think for the publicly available ones I don't worry about that [privacy] because they've gotten permission to be used. I think probably the thing I try to emphasize with my students is to try to not be too overly critical of what they see because these are teachers who have opened up their classrooms and we have to honor that they have done that' (Interviewee 24).

Interviewee 38 explained that concerns about potential harms extend to the handling and reuse of the data throughout its lifecycle, including the use of video for classroom instruction. Teaching students how to responsibly produce, share, and reuse video records of practice includes practice using a secure online video sharing platform, as well as training about how to responsibly reuse video in a classroom environment:

> 'We ask our students to take the video and upload it to a secure server space that we have posted here at [university]. We only access the video through that secure server space or through Edthena, which you need a password to get on. We also, in terms of its use in class, I think this holds true for both the video of other people teaching and the video of our student teaching, is

> we really push against tendencies to be evaluative or judgmental. Comments like, "Oh, that's bad," or, "I'd never do that that way," are discouraged explicitly and we really try to dig into what we see happening. We focus on teacher talk, on the student talk, on tasks and tools that are used. We really have an emphasis throughout our program on what we call leveraging student thinking' (Interviewee 38).

For these interviewees, responsible data practices are part of educating future teachers and researchers. Emphasizing that data reusers and producers put themselves at risk when they put their subjects at risk is one way to educate students about working with video records of practice. Interviewees 44, 24, and 38 previously quoted, used this strategy that emphasized appropriate versus inappropriate data sharing methods because sharing video records of practice was a significant part of their teaching. Other interviewees focused on placing limitations on data sharing, rather than on providing education and training for data handling and use.

Informed consent: Facilitating and/or restricting data sharing

Participants discussed institutional review board (IRB) or ethics panel decisions and informed consent as tools that could be used to guide, limit, or prevent sharing and reuse of qualitative data, including video records of practice. In addition to protecting the privacy of individual research subjects, participants were concerned with confidentiality and the mediation of access based on perceived risks of information disclosure.

Interviewee 2 explained that as a data producer, she has purposely written restrictive informed consent forms that would limit sharing and reuse, as a way to protect the subjects represented in the data:

> 'I'm very protective of my video and audio because of how I wrote my consent forms. That people were not going to be watching that video who weren't part of the research team. And I wrote it that way because I didn't know if I was going to invite someone to use my data later with me. But I just saw my custodial role as someone who needed to protect identities in that video as well' (Interviewee 2).

Informed consent was important for data reuse among our interviewees, who said that it was critical to understand the parameters of the original research and what subjects had agreed to before reusing data for a new purpose. For example, Interviewee 2 explained that even when qualitative data is available from the data producer or a repository, it is important for reusers to understand the informed consent agreements, in order to understand whether subjects agreed to their data being shared beyond the original data producer and whether their consent was for the data to be used only for specific purposes. She argued that as a reuser, she would want to see the approved IRB application and/or consent forms in order to understand whether she could ethically reuse the data for her purposes:

> 'I actually wanted to see the original IRB, 'cause I wanted to know what was consented. If you're looking at video data and the consent was entirely written around teacher behavior, I don't know if ethically you can look at students' behavior or actions or activity. I think that that is actually a big problem and that if you looked at APA [American Psychological

> Association] ethics standards, there actually could be some really serious problems with that, and so that's a big question I have around reuse. I would want to see the IRB' (Interviewee 2).

Interviewee 34 also discussed selecting data for reuse from the perspective of the original data producer. He explained that older consent forms did not include the option for participants to choose whether their images would be shared widely. This raises the question of whether consent forms that lack an explicit statement one way or another regarding data sharing should be assumed to allow or prohibit sharing. The Economic and Social Research Council (ESRC) Qualitative Data Archival Resource Centre in the UK has recommended that data producers obtain permission whenever possible, clearly stating what they intend to do with the data (Corti, Day and Backhouse, 2000), as a means of narrowing the scope of data sharing to mitigate risk and reduce potential harms. The debate about whether data sharing should depend on opt-in or opt-out statements on consent forms arose in several interviews, and our participants were divided about whether data should be reused only for purposes that subjects explicitly agreed to, or whether data could be reused for any purpose that is not expressly prohibited:

> 'When it comes to the selection it's all a matter of, do you have the permissions or do you not have the permission? Currently, in our consent forms, our older ones didn't have this detail. The current ones, people can select. I know you can use this for research only, or you could use it to present at conferences or you could do it to generate publishable material' (Interviewee 34).

Respecting the privacy of the subjects by not sharing data without explicit permission was described as both a necessary protection and also an impediment to research and teaching. For Interviewee 29, videos collected for research cannot be shared or used for other purposes, but videos that are publicly available are considered appropriate data sources to use for training new research assistants:

> 'It was really important to be able to use data that was publicly available to do this sort of thing [training new coders on a large research project]. Some of our internal videos we have that type of permission from that. I think I couldn't really use any of the videos that I'd collected through research, because I don't have that similar sort of public permission in that case' (Interviewee 29).

This not only limits sharing data for future research uses, but also the use of the video for teacher professional development, "because I have videos that I have collected for research purposes and I strategically don't show them because I don't have permission to show them in that capacity. Sure, I definitely want to make sure that I follow privacy guidelines" (Interviewee 27).

For both in-service and pre-service teachers who record themselves teaching, consent is typically required from the parents of their students if students are recorded. Interviewee 38 described a process in which parents are given the opportunity to opt in or out from having video that includes their children be used for teacher education, research, academic conferences, publications, and future professional uses for the teacher in the video such as job applications:

> '[W]hen it's our teachers recording themselves, but then in terms of using it, even though we have consent forms that give the guardians and the young people a chance to indicate through their signature and check boxes how that video might be used. The ways that we give them – the indicators that we ask them to check or choose to not check are that the video will be used for class discussion and tasks, in class learning tasks, in our teacher education program. They will be used for program development purposes within our program and department. They'll be used to present and disseminate findings of research at, and these are IRB approved forms, at professional conferences and potentially through journal articles or other academic outlets. We have another box that parents and young people can check that will allow our teacher candidates to use clips from that video for job application purposes. They can choose to check or not check any of those boxes. If we have someone who doesn't check one of those boxes and their face appears on the video that video is deleted or that segment of the video is clipped' (Interviewee 38).

Interviewee 39 also discussed concerns about data sharing and reuse with regard to the privacy of subjects represented in the data. She argued that even when subjects are given the opportunity to opt in or out, it is very difficult to convey the risks of participating as a research subject. She also argued that ethical research includes permitting subjects to withdraw from the study if they do not feel comfortable with their portrayal in the final product. In her opinion, repositories are following the letter of the law but are not taking all of the steps required for truly ethical data practices that protect the subjects represented in video data:

> 'I don't know that we think through the power of images and the challenges of considering the future use of video when we create video. I think that's just a challenge inherent to the medium itself, and I don't know that organizations think through that. That's what to me is the most difficult problem to solve, that you have to really think through about what you're putting online and how it can be used in the future. Some of it will always be out of your control and so that's a big problem. You can do some things, you can be very careful about disclosing all these risks to the people that you are videotaping so that they understand how these videos will be shared with others, what information about them and the videos will be shared and whatnot, and you give them the possibility to see the final product and withdraw from the project if they don't feel comfortable with the way they're portrayed. I don't know that organizations are following all these ethical steps versus just having a legal form signed that maybe lists a few risks but it doesn't necessarily give a full picture of what the final product really is going to look like. If I were the one in charge, I would have maybe more constraints to those who create videos to make sure that those who are portrayed in the videos are highly protected' (Interviewee 39).

Interviewee 41 also stressed the responsibility of data producers with regard to future use of qualitative data, "I definitely think it's important to be careful and conscientious of the student identities and make sure that your video is being used in the way that it's stipulated in your consent form and your IRB. I think there are ways that it can be misused" (Interviewee 41).

This sentiment, that appropriate reuse of data includes not only doing what is technically and legally permissible but also doing so in a way that is respectful and protective of the subjects represented in the data, was shared by several participants. Participant 41 described research that focuses on the capabilities or motivations of students as problematic, "It could become problematic if one were writing a piece that was not approved or reviewed [by the IRB] about a student's motivation or capability in this video" (Interviewee 41). The need to protect the subjects in qualitative data extends to teachers as well as students. Interviewee 34 explained that he would feel uncomfortable sharing video that shows a teacher in a poor light: "if you have one that you feel is a little incriminating of a teacher and their practice, I wouldn't share it" (Interviewee 34).

This belief, that the best way to protect subjects is to restrict or limit data sharing even when they have provided their informed consent, was a common theme among our interviewees. Among interviewees who discussed limiting data sharing, one common method that arose was to restrict what was captured in the data production process.

Limiting data capture/production

While anonymization remains a significant problem for qualitative data even if research subjects have given their consent for the data to be shared over time, researchers do use some anonymization measures, including limiting the scope of the video data captured. Considering video records of practice, one interviewee stated: "We do try to maintain the anonymity wherever we can, but it's just not really possible when you're talking about video" (Interviewee 15).

For data producers, such as Interviewee 15, anonymization is challenging because strategically capturing video and/or audio around particular individuals in a dynamic and sometimes unpredictable classroom environment is difficult. In many cases, anonymization is an activity that must happen after the data collection event:

> 'If we have someone who doesn't check one of those boxes and their face appears on the video, that video is deleted or that segment of the video is clipped' (Interviewee 38).

For participants who produce their own data, limiting the scope of what was captured by the camera is one way to protect the subjects in the video. "Well for example when I did my study I didn't include any student faces. So it's only showing their hands, and you can hear their voice. But you could never see their faces" (Interviewee 16). While this does not guarantee the anonymity of the students in the video, it significantly reduces the risk of future privacy violations by either the original data producer or potential reusers of the data.

Discussion

In summary, we found that data producers and reusers were concerned about the risks that qualitative data, and video records of practice in particular, present to themselves, their colleagues, and the subjects represented in the data. Specifically, they emphasized risks relating to the privacy of the teachers and students who appeared in the videos. In response to these risks, interviewees engaged in a number of strategies to minimize risk

and/or mitigate harm including: (1) education and training; (2) using informed consent to facilitate and/or restrict data sharing; and (3) limiting data capture/production.

This discussion section focuses on our two research questions:

- What concerns do data producers and reusers have regarding the privacy and confidentiality of subjects seen in video records of practice and how are these concerns represented in the risks or potential harms they perceive?
- How do perceptions of risk or potential harm influence attitudes about data sharing and reuse among researchers and teacher-educators in the field of education?

We have reported here on the results of 44 interviews with researchers and teacher-educators who reuse digital records of practice for research and/or teaching. Our findings demonstrate varying levels of comfort with data sharing and reuse, and raise questions for us about who should be responsible for ensuring that publicly available video records of practice remain accessible.

Perceptions of Harm

Our findings provide a more nuanced understanding of how data producers and reusers conceive of harm in personally identifiable data than has been present in information science research about data sharing and reuse to date. Previous studies have focused on ways to facilitate data sharing and reuse while following the restrictions set by consent forms and IRB determinations, rather than on the underlying privacy concerns of data producers, reusers, and subjects (e.g. Corti, Day and Backhouse, 2000; Lagoze et al., 2013). In this article, we have found that interviewees view privacy concerns around the sharing and reuse of qualitative data in education, especially video records of practice, as a double-edged sword of reputational harm, in which data producers and reusers, as well as the subjects represented in the data face risks. The close ties between data producers, reusers, and subjects in the field of education heighten these risks because they make it more difficult for subjects to remain anonymous.

This article links information science research, which has focused on the practical and technical aspects of data sharing and reuse (e.g. Lagoze et al., 2013) and has tended to treat ethical considerations as challenges to be overcome (Corti, Day and Backhouse, 2000), with disciplinary research from fields that have long histories of working with human subjects such as medicine and education (e.g. El Emam and Moher, 2013; Hammersley, 2015).

Our interviewees talked about the negative consequences for teachers and students that they thought could result from sharing video in which individuals are identifiable. In addition to the risk that those individuals would be singled-out based on their appearance or behaviors in records of practice, interviewees were concerned that such an event would also have negative consequences for the data producer who made the data available. Researchers in the field of information science, such as Yoon (2016), have discussed harm to data producers in relation to the quality of the data that they share. We argue here that data producers and reusers in the field of education are also concerned with the potential harm that may occur to themselves and their subjects as a result of how the data content is reused and discussed, and that these concerns also influence their attitudes about data sharing and reuse.

Attitudes about Data Sharing and Reuse

Many authors have discussed the reluctance of members of other disciplines to share data, disciplinary members' concerns include bring scooped (Wallis, 2014), the difficulty of preparing data for sharing (Tenopir et al., 2011), and the lack of a good mechanism for data sharing (Alter and Vardigan, 2015). In the field of education our findings suggest that data sharing and reuse are seen as being in conflict with disciplinary norms around responsible and ethical conduct for both researchers and teacher-educators. Data producers described protecting the students and teachers who appear in their data as an important responsibility and an extension of their role as educators. This raises questions about who should be responsible for protecting the subjects who appear in video records of practice.

Interviewees were confident in their own ability to protect the subjects in their data, but skeptical of both the motivations and abilities of others. They were also skeptical about whether the subjects who consented to appear in their data truly understood the risks that they were taking. This skepticism about informed consent is consistent with previous findings from research about privacy and human subject research (e.g. Manhas et al., 2015; Moser et al., 2005; Sin, 2005).

Interviewees also expressed doubt about the legitimacy of publicly available data in repositories, specifically about whether the data producers really had permission for the data to be publicly shared based on informed consent forms and IRB determinations. Questions about whether data sharing should depend on explicit permission to share, or simply a lack of prohibition against sharing, were a reflection of the tension that our interviewees experienced between data sharing and wanting to protect themselves and their subjects. These concerns also emerged in conversations about how interviewees controlled access to the video records of practice that they produced.

Controlling Access to Data

One of the ways that our interviewees responded to risks concerning the privacy of the subjects in their data was by taking a strict stance toward data capture, processing, and management in order to exercise tight control over access and reuse. Interviewees reported using informed consent forms and IRB applications as tools to control access to, and reuse of, video records of practice. By writing consent forms and IRB documents that expressly allow or prohibit data sharing and/or reuse, data producers leverage the tools at their disposal to protect the subjects represented in their data by preventing others from having access.

Discussions with our interviewees about informed consent included debates about whether permission must be expressly *given* for data sharing, or whether data sharing must be specifically *prohibited*. Participants described both leveraging the lack of prohibition in order to share and reuse data, as well as including language prohibiting sharing or reuse as a way to avoid having to provide access to their data for others. In some cases, vague permission statements were interpreted broadly by producers and reusers to facilitate data sharing, and in others they were interpreted narrowly in order to protect the data producers and subjects by preventing sharing and reuse.

This issue has been explored by researchers, such as Corti, Day and Backhouse (2000), who recommended that explicit permission should be obtained whenever possible for actions, such as depositing data into a repository for preservation or sharing. We found that the data producers we interviewed generally agreed that data should not be shared without the express permission of subjects. Data reusers were more

likely to assume that if data was available to them, then those permissions had been given. Reusers often did not take any steps to check, preferring to ask for forgiveness rather than permission in case those permissions would prevent or limit their own reuse. Repositories can support both data producers and reusers by helping to make permissions and/or restrictions for data visible at the point of access.

Conclusion

This study found that data producers and reusers in the field of education view data sharing and reuse as fraught activities in light of the potential harm to data producers, reusers, and subjects that our interviewees discussed. Privacy concerns that influence attitudes about data sharing and reuse among educational researchers and teacher-educators echo those that have been expressed by data users in other fields (e.g. Borgman, 2012; Custers and Uršič, 2013; Lagoze et al., 2013) but are intensified by the unique nature of video records of practice in education and the challenges that they present with regard to privacy. Data producers and reusers in education use the tools at their disposal, such as informed consent and the IRB process, to protect their research subjects, as well as themselves, by restricting access to their data. Efforts to facilitate sharing and reuse of video records of practice in education should take these unique concerns into account, as well as the strong desire of data producers and reusers to protect both their subjects as well as themselves.

Acknowledgements

This research was made possible in part by the Institute of Museum and Library Services, National Leadership Grant # LG-06-14-0122-14: “Qualitative Data Reuse: Records of Practice in Educational Research and Teacher Development.”

References

Alter, G.C. & Vardigan, M. (2015). Addressing global data sharing challenges. *Journal of Empirical Research on Human Research Ethics*, *10*(3), 317–323. doi:10.1177/1556264615591561

Andersson, E. & Sørvik, G.O. (2013). Reality lost? Re-use of qualitative data in classroom video studies. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *14*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1941

Arafeh, S. & McLaughlin, M. (2002). *Legal and ethical issues in the use of video in education research*. National Center for Education Statistics. Retrieved from http://nces.ed.gov/pubs2002/200201.pdf

Arzberger, P.W., Schroeder, P., Beaulieu, A., Bowker, G.C., Casey, K., Laaksonen, L., … Wouters, P. (2004). Promoting access to public research data for scientific, economic, and social development. *Data Science Journal*, *3*(29), 135–152. doi:10.2481/dsj.3.135

Association of Research Libraries. (n.d.). White House directive on public access to federally funded research and data. Retrieved from http://www.arl.org/focus-areas/public-access-policies/federally-funded-research/2696-white-house-directive-on-public-access-to-federally-funded-research-and-data

Bass, H., Usiskin, Z., Burrill, G., National Research Council (U.S.), Mathematical Sciences Education Board, & United States National Commission on Mathematics Instruction (Eds.). (2002). *Studying classroom teaching as a medium for professional development: Proceedings of a U.S.-Japan workshop*. Washington, DC: National Academy Press. Retrieved from http://www.nap.edu/catalog/10289.html

Berg, M. & Goorman, E. (1999). The contextual nature of medical information. *International Journal of Medical Informatics*, *56*(1–3), 51–60. doi:10.1016/S1386-5056(99)00041-6

Bishoff, C. & Johnston, L. (2015). Approaches to data sharing: An analysis of NSF data management plans from a large research university. *Journal of Librarianship and Scholarly Communication*, *3*(2). doi:10.7710/2162-3309.1231

Bishop, L. (2009). Ethical sharing and reuse of qualitative data. *Australian Journal of Social Issues*, *44*(3), 255–272. doi:10.1002/j.1839-4655.2009.tb00145.x

Blomberg, G., Stürmer, K., & Seidel, T. (2011). How pre-service teachers observe teaching on video: Effects of viewers' teaching subjects and the subject of the video. *Teaching and Teacher Education*, *27*(7), 1131–1140. doi:10.1016/j.tate.2011.04.008

Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the Association for Information Science and Technology*, *63*(6), 1059–1078. doi:10.1002/asi.22634

Burleigh, J.C., & Peterson, H.W. (1967). Videotapes in teacher education. *The Elementary School Journal*, *68*(1), 35–38.

Carlson, S. & Anderson, B. (2007). What are data? The many kinds of data and their implications for data re-use. *Journal of Computer-Mediated Communication*, *12*(2), 635–651. doi:10.1111/j.1083-6101.2007.00342.x

Cheville, R.A. (2016). Data sharing from a policy perspective. *Advances in Engineering Education*, *5*(2), 17.

Cohen, M., Burr, V., Goetz-Haver, S., & Morales, M. (2003). Digital video and teacher education. Paper presented at the Society for Information Technology & Teacher Education International Conference. Retrieved from https://www.learntechlib.org/p/18184/

Corti, L. (2000). Progress and problems of preserving and providing access to qualitative data for social research – The international picture of an emerging culture. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, *1*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1019

Corti, L. (2007). Re-using archived qualitative data – Where, how, why? *Archival Science*, *7*(1), 37–54. doi:10.1007/s10502-006-9038-y

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). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/viewArticle/1024

Custers, B. & Uršič, H. (2016). Big data and data reuse: A taxonomy of data reuse for balancing big data benefits and personal data protection. *International Data Privacy Law*, *6*(1), 4–15. doi:10.1093/idpl/ipv028

Darian-Smith, K. & Henningham, N. (2012). Social research and the privacy and participation of children: Reflections on researching Australian children's playlore. *Children & Society*, *28*(4), 327–338. doi:10.1111/j.1099-0860.2012.00475.x

Derry, S.J., 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. doi:10.1080/10508400903452884

Doyle, T. & Veranas, J. (2014). Public anonymity and the connected world. *Ethics and Information Technology*, *16*(3), 207–218. doi:10.1007/s10676-014-9346-5

El Emam, K. & Moher, E. (2013). Privacy and anonymity challenges when collecting data for public health purposes. *Journal of Law, Medicine & Ethics*, *41*(1_Suppl), 37–41. doi:10.1111/jlme.12036

Eschenfelder, K. & Johnson, A. (2011). The limits of sharing: controlled data collections. *Proceedings of the American Society for Information Science and Technology*, *1*(48), 1–10. doi:10.1002/meet.2011.14504801062

Faniel, I.M. & Jacobsen, T.E. (2010). Reusing scientific data: How earthquake engineering researchers assess the reusability of colleagues' data. *Computer Supported Cooperative Work (CSCW)*, *19*(3–4), 355–375. doi:10.1007/s10606-010-9117-8

Flewitt, R. (2006). Using video to investigate preschool classroom interaction: Education research assumptions and methodological practices. *Visual Communication*, *5*(1), 25–50. doi:10.1177/1470357206060917

Frank, R.D., Kriesberg, A., Yakel, E., & Faniel, I.M. (2015). Looting hoards of gold and poaching spotted owls: Data confidentiality among archaeologists and zoologists. In *Proceedings of the 78th ASIS&T Annual Meeting: Information Science with Impact: Research in and for the Community* (pp. 37:1–37:10). Silver Springs, MD, USA: American Society for Information Science. Retrieved from http://dl.acm.org/citation.cfm?id=2857070.2857107

Frank, R.D., Suzuka, K., & Yakel, E. (2016). Examining the reuse of qualitative research data: Digital video in education. In *Archiving Conference* (Vol. 2016, pp. 146–151). Washington, DC: Society for Imaging Science and Technology. doi:10.2352/issn.2168-3204.2016.1.0.146

Gilmore, R.O., Adolph, K.E., Millman, D.S., & Gordon, A. (2016). Transforming education research through open video data sharing. *Advances in Engineering Education*, *5*(2), 17. Retrieved from http://advances.asee.org/wp-content/uploads/vol05/issue02/Papers/AEE-18-Gilmore.pdf

Hammersley, M. (2015). Research ethics and the concept of children's rights. *Children & Society*, *29*(6), 569–582. doi:10.1111/chso.12077

Holdren, J.P. (2013). *Increasing access to the results of federally funded scientific research* (Memorandum). Washington, D.C.: Office of Science and Technology Policy. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf

Holsti, O.R. (1969). *Content analysis for the social sciences and humanities*. Reading, MA: Addison-Wesley Publishing Company.

Inter-University Consortium for Political and Social Research. (n.d.). Confidentiality. Retrieved from https://www.icpsr.umich.edu/icpsrweb/content/datamanagement/confidentiality/index.html

Joyce, N. (2011). An analysis of the extent of the juvenile offender's right to privacy: Is the child's right to privacy circumvented by public interest? *European Journal of Crime, Criminal Law and Criminal Justice*, *19*(2), 113–124. doi:10.1163/157181711X566335

Lagoze, C., Block, W.C., Williams, J., Abowd, J., & Vilhuber, L. (2013). Data management of confidential data. *International Journal of Digital Curation*, *8*(1), 265–278. doi:10.2218/ijdc.v8i1.259

Manhas, K.P., Page, S., Dodd, S.X., Letourneau, N., Ambrose, A., Cui, X., & Tough, S. C. (2015). Parent perspectives on privacy and governance for a pediatric repository of non-biological, research data. *Journal of Empirical Research on Human Research Ethics*, *10*(1), 88–99. doi:10.1177/1556264614564970

Marx, G.T. (1999). What's in a name? Reflections on the sociology of anonymity. *The Information Society*, *15*, 99–112. doi:10.1080/019722499128565

Masats, D. & Dooly, M. (2011). Rethinking the use of video in teacher education: A holistic approach. *Teaching and Teacher Education*, *27*(7), 1151–1162. doi:10.1016/j.tate.2011.04.004

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook*. Thousand Oaks: Sage Publications.

Mondada, L. (2014). Ethics in action: Anonymization as a participant's concern and a participant's practice. *Human Studies*, *37*(2), 179–209. doi:10.1007/s10746-013-9286-9

Moore, N. (2007). (Re)Using qualitative data? *Sociological Research Online*, *12*(3). doi:10.5153/sro.1496

Moser, C., Chen, T., & Schoenebeck, S. (2017). Parents' and children's preferences about parents sharing about children on social media. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '17)*. Denver, CO. doi:10.1145/3025453.3025587

Narayanan, A. & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *Proceedings of the 2008 Symposium on Security and Privacy* (pp. 111–125). IEEE. doi:10.1109/SP.2008.33

Penn-Edwards, S. (2012). Human factors affecting the use of video recording methodology in qualitative research. *International Journal of Multiple Research Approaches*, *6*(2), 150–159. doi:10.5172/mra.2012.6.2.150

Privacy. (2008). In L. Given, *The SAGE Encyclopedia of Qualitative Research Methods*. Thousand Oaks, C.A.: SAGE Publications, Inc.

Rook, M.M., & McDonald, S.P. (2012). Digital records of practice: A literature review of video analysis in teacher practice. Presented at the Society for Information Technology & Teacher Education International Conference. Retrieved from http://www.editlib.org/p/39784/

Saldaña, J. (2015). *Thinking qualitatively: Methods of mind*. Thousand Oaks, California: SAGE.

Sin, C.H. (2005). Seeking informed consent: Reflections on research practice. *Sociology*, *39*(2), 277–294. doi:10.1177/0038038505050539

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), e21101. doi:10.1371/journal.pone.0021101

van Es, E. & Sherin, M. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596.

van House, N. (2002). Digital libraries and practices of trust: Networked biodiversity information. *Social Epistemology: A Journal of Knowledge, Culture and Policy*, *16*(1), 99–114. doi:10.1080/02691720210132833

Wallace, K.A. (2008). Online anonymity. In K. E. Himma & H. T. Tavani (Eds.), *The handbook of information and computer ethics* (pp. 165–189). Hoboken, NJ, USA: John Wiley & Sons, Inc.

Wallis, J. (2014). Data producers courting data reusers: Two cases from modeling communities. *International Journal of Digital Curation*, *9*(1), 98–109. doi:10.2218/ijdc.v9i1.304

Yakel, E., Suzuka, K., & Frank, R.D. (2018). Qualitative data reuse: Records of practice in educational research and teacher development – Interview protocol and qualitative data codeset for the interviews. Ann Arbor, MI: DeepBlue Data. doi:10.7302/Z28C9TGP

Yardley, S.J., Watts, K.M., Pearson, J., & Richardson, J.C. (2013). Ethical issues in the reuse of qualitative data: Perspectives from literature, practice, and participants. *Qualitative Health Research*, *24*(1), 102–113. doi:10.1177/1049732313518373

Yoon, A. (2016). Data reusers' trust development. *Journal of the Association for Information Science and Technology*, *68*(4), 946–956. doi:10.1002/asi.23730

Zhang, M., Lundeberg, M., Koehler, M. J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education*, *27*(2), 454–462. doi:10.1016/j.tate.2010.09.015

Zimmerman, A. (2007). Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries*, *7*(1–2), 5–16. doi:10.1007/s00799-007-0015-8

Zimmerman, A. (2008). New knowledge from old data: The role of standards in the sharing and reuse of ecological data. *Science, Technology & Human Values*, *33*(5), 631–652. doi:10.1177/0162243907306704