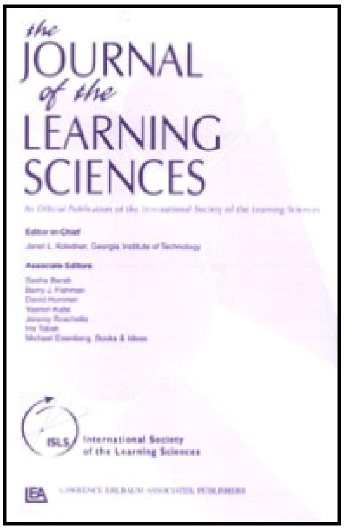

Journal of the Learning Sciences
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653672

Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics

Sharon J. Derry [a]; Roy D. Pea [b]; Brigid Barron [b]; Randi A. Engle [c]; Frederick Erickson [d]; Ricki Goldman [e]; Rogers Hall [f]; Timothy Koschmann [g]; Jay L. Lemke [h]; Miriam Gamoran Sherin [i];Bruce L. Sherin [i]
[a] Department of Educational Psychology, University of Wisconsin-Madison, [b] School of Education, Stanford University, [c] Graduate School of Education, University of California Berkeley, [d] Department of Education, University of California Los Angeles, [e] Department of Administration, Leadership and Technology, New York University, [f] Department of Teaching and Learning, Vanderbilt University, [g] Department of Medical Eduation, Southern Illinois University, [h] Department of Educational Studies, University of Michigan, [i] School of Education and Social Policy, Northwestern University,

Online publication date: 25 January 2010

To cite this Article Derry, Sharon J. , Pea, Roy D. , Barron, Brigid , Engle, Randi A. , Erickson, Frederick , Goldman, Ricki , Hall, Rogers , Koschmann, Timothy , Lemke, Jay L. , Sherin, Miriam Gamoran andSherin, Bruce 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

To link to this Article: DOI: 10.1080/10508400903452884

URL: http://dx.doi.org/10.1080/10508400903452884

THE JOURNAL OF THE LEARNING SCIENCES, 19: 3–53, 2010

ISSN: 1050-8406 print / 1532-7809 online
DOI: 10.1080/10508400903452884

Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics

Sharon J. Derry
Department of Educational Psychology
University of Wisconsin–Madison

Roy D. Pea and Brigid Barron
School of Education
Stanford University

Randi A. Engle
Graduate School of Education
University of California Berkeley

Frederick Erickson
Department of Education
University of California Los Angeles

Ricki Goldman
Department of Administration, Leadership and Technology
New York University

Rogers Hall
Department of Teaching and Learning
Vanderbilt University

Timothy Koschmann
Department of Medical Eduation
Southern Illinois University

Jay L. Lemke
Department of Educational Studies
University of Michigan

Miriam Gamoran Sherin and Bruce L. Sherin
School of Education and Social Policy
Northwestern University

Correspondence should be addressed to Sharon J. Derry, University of Wisconsin–Madison, Department of Educational Psychology, 1025 West Johnson Street, Madison, WI 53706. E-mail: sjderry@education.wisc.edu

Focusing on expanding technical capabilities and new collaborative possibilities, we address 4 challenges for scientists who collect and use video records to conduct research in and on complex learning environments: (a) *Selection:* How can researchers be systematic in deciding which elements of a complex environment or extensive video corpus to select for study? (b) *Analysis:* What analytical frameworks and practices are appropriate for given research problems? (c) *Technology:* What technologies are available and what new tools must be developed to support collecting, archiving, analyzing, reporting, and collaboratively sharing video? and (d) *Ethics:* How can research protocols encourage broad video sharing and reuse while adequately protecting the rights of research participants who are recorded?

> Every technology has a philosophy which is given expression in how the technology makes people use their minds, in what it makes us do with our bodies, in how it codifies the world, in which of our senses it amplifies, in which of our emotional and intellectual tendencies it disregards. (Neil Postman, as cited in Gore, 2007, p. 20)

Neil Postman's proclamation aptly fits the case of video research in the learning sciences. Rapid development and widespread availability of affordable, usable, high-quality video technology is transforming the practice of learning science research. Because new video technologies provide powerful ways of collecting, sharing, studying, presenting, and archiving detailed cases of practice to support teaching, learning, and intensive study of those practices, many learning science research projects now incorporate a substantial video component.

The goal of this contribution is to provide guidance for researchers conducting (a) studies that use video to closely examine teaching and learning in learning environments such as classrooms, the goals of which are to understand learning processes and better design formal learning environments; and (b) studies using video for in-depth analyses of peer and adult–child interactions in informal settings, such as museums and homes, the goals of which are to help researchers and developers understand informal learning as it occurs naturally in various contexts and to achieve better informal settings for learning. We collaborated to provide a careful discussion of principles, strategies, and important issues that should inform researchers' choices at all points in the video research process.[1] We tried to develop

[1]Motivated by an increasing number of research proposals that request funding for collecting and analyzing video, in December 2005 the National Science Foundation sought guidance from qualified researchers to help address the question "What does good video research look like?" An interdisciplinary conference of scholars who conduct educational research with video convened with the goal of providing guidance for understanding and judging the video research proposed to and funded by programs within the National Science Foundation's Directorate for Education and Human Resources. The authors of this article are a subset of scholars who continued to meet in person and online following the conference to vet and synthesize its major findings and summarize them as guidelines for members of the learning science community.

these guidelines *without favoring any particular methodological orientation or set of methods*. The learning sciences is an interdisciplinary field, and video is a tool that enhances various methodologies associated with different, and some would argue incommensurate, philosophical orientations. These include ethnography, ethnomethodology, experimentation, discourse analysis, interaction analysis, and others. But regardless of a researcher's methodological orientation or specific research goals, video offers a means of close documentation and observation and presents unprecedented analytical, collaborative, and archival possibilities, as well as new problems. Researchers with different methodological orientations and goals confront many common challenges among those that we address.

Our work was a search for *boundary objects* in the sense defined by Star (1989), who considered how "like the blackboard, a boundary object sits in the middle of a group of actors with divergent viewpoints" (p. 46). Bowker and Star (2000) argued that boundary objects can be ideas, tools, stories, memories, and material items that circulate in networks of actors across situations, playing important roles in communication and knowledge building within and across communities. Here we present a discussion of some boundary objects that seem particularly worthy of attention from all video researchers in the learning sciences regardless of disciplinary or methodological orientation.

To encourage creativity as well as open our community to broad interdisciplinary understandings of how video records can be used effectively to collaboratively build knowledge about human activities, we advocate maintaining diversity and flexibility in these boundary objects. At the same time, we recognize that progress as a research community will depend on moving toward standardization and widespread use of agreed-upon boundary objects, because our ability to communicate and share across research groups requires some degree of commonality among practices and tools (Derry, 2007b). Consider the aphorism *Words are the chains that set us free.* This paradoxical concept makes more sense, Wittgenstein (1953) explained in his *Philosophical Investigations*, when considered in light of the problem of other minds. Wittgenstein helped us understand that words set us free in the generative sense that each new occasion of language use creates a unique expression, different in at least some respects from any other occasion in terms of context, speaker history, and intent. But these words are also chains because there must be some commonality, a constraint in meaning, to provide the very possibility of building bridges between minds, of sharing ideas and perspectives that another mind can understand. We draw an analogy between words used for communication and standard boundary objects that support scientific knowledge building and communication through video research. Such boundary objects will be essential in order for the learning science community to build a cumulative scientific knowledge base from working with video. Yet questions arise regarding what kinds of boundary objects can be agreed to with little controversy, and what attempts at standardization are likely to mire the community in debate or stifle innovation.

This contribution focuses on boundary objects related to the collection and use of video records to produce data for research in and on complex learning environments.[2] We address four sets of challenges:

1. *Selection.* How does a researcher decide which elements of a complex environment should be recorded, or which aspects of an extensive video corpus should be sampled for further examination?
2. *Analysis.* What analytical frameworks and practices are available, and which of these are scientifically valid and appropriate for given research problems?
3. *Technology.* What technological tools are available and which social tools must be developed and disseminated to support collecting, archiving, analyzing, reporting, and collaboratively sharing video?
4. *Ethics.* How can research protocols encourage broad video sharing and reuse while adequately protecting the rights of the human subjects who are represented in such recordings?

Our discussion builds on previous methodological treatments (e.g., Chi, 1997; Jordan & Henderson, 1995; Schatzman & Strauss, 1973) but extends those analyses by raising interesting new issues that surface in consideration of expanding technical capabilities and collaborative possibilities. We hope our effort not only informs researchers and funding agencies associated with the learning sciences but also encourages and facilitates data sharing and collaboration for knowledge building across projects; speeds the learning curve for new researchers; and helps ensure that time, effort, and scarce resources are expended wisely.

SELECTION IN VIDEO RESEARCH

Accessible video technologies provide researchers with powerful "microscopes" that greatly increase the interactional detail that can be obtained and permanently stored for comprehensive analysis and reanalysis by multiple investigators. However, this enhanced observational power requires thoughtful attention to the problem of how to extract data and meaning from the large, complex video corpora that such research creates. This is one case of *data selection*—a process of focusing on particular information in accordance with the theoretical frameworks, research questions, and instruments a researcher chooses. What follows is an analysis of the *selection problem* viewed as a boundary concept for research in which video is a

[2]Discussion of creating and using video for teaching, learning, and assessment is omitted in this presentation, but an excellent discussion can be obtained from the original report upon which this article is based (Sherin & Sherin, 2007).

major (although almost never the only) data source. The question we address is how to be systematic rather than arbitrary when selecting with video recording devices and from the video recordings they create.

Conceptualizing the Nature of What Is Selected From Video

We begin by asking what entity is created when particular elements from complex environments are captured by video, removed from their larger contexts, and eventually become "clips" for further examination. There are various ways in which to conceptualize the video clip, but one useful perspective from perceptual psychology (Zacks & Tversky, 2001) is that video segments represent *events*. Any video corpus captures many events. Selection determines which events are brought into focus for deeper analysis.

Events are time-analogs of objects. Like objects, they have underlying structures reflecting multiple parts and timescales (Lemke, 2000). Consider a classroom event during which students collaborate on the interpretation of data presented as a histogram. This event could be parsed in terms of various subevents: for example, the presentation of an idea by a student; the response by another indicating acceptance, confusion, disagreement, or even disengagement; periods of negotiation through which joint understanding emerges; and so on. These subevents might be further analyzed as coordination of even smaller events on smaller timescales: gestures, speech, tool use, mental states, and so on. Or the entire classroom event could be considered to be part of a longer term macroevent, such as the development of students' ability to read and interpret a range of data representations from various sources.

The ability to decompose a complex event and select specific parts to pay further attention to is jointly influenced by a researcher's perception and by what actually occurs in the presence of the recording device (Goldman-Segall, 1998; Leacock, 1973). Psychological studies have shown that people often "see" events similarly in terms of causal, behavioral, and thematic structures, although *professional vision* (Goodwin, 1995; see also "disciplined perception," Stevens & Hall, 1998), or expert ways of interpreting events, develops through specialized training and experience. Observed changes in types of behavior (e.g., laughing vs. talking), physical direction (turning toward the exit), the object of behavior (child puts down pencil, picks up toy), setting (in the dining room vs. in the kitchen), and tempo of activity (jogging vs. cool-down phase) determine how people see and interpret events as chunks of time. The trained observer can reasonably conclude, "The theme of this video chunk is doing math, I'm interested in that; that one is mostly about *play*, I'll discard that chunk. However, this one is both *play and doing math*, and so I'll select it for further study."

The researcher's specific interest will determine which events and which timescales a study should select. For example, if the researcher is interested in studying the coordination of the cognitive, material, and social processes through which understanding of the concept *variable* first arises, an appropriate plan would be to begin by recording, or by identifying in an appropriate set of previously recorded data, problem-solving situations in which children encounter this concept when the concept is first developing. A reasonable timescale for each selection would be the period in which a child encounters, grapples with, and resolves a problem involving variables. However, if the interest is in studying the development of how students acquire a more inclusive concept (such as the histogram) as a representational tool, then the researcher must focus on a larger event, including antecedents occurring over a longer timescale. A reasonable selection strategy in this case might consist of identifying important "critical events" (Powell, Francisco, & Maher, 2003, p. 413) within a larger developmental time frame.

Selection at Different Stages of Inquiry

Although video research is not a linear progression of steps (e.g., Cobb & Whitenack, 1996; Pea & Hoffert, 2007; Powell et al., 2003), the inquiry process can be conceptualized as moving among a number of phases: (a) planning a study, (b) shooting original footage, (c) choosing one or more clips from a corpus of such footage, (d) focusing on the selected video clip or clips in appropriate ways depending on the researcher's goals, (e) developing final products for presentation, and (f) addressing issues related to archiving and curatorship of video and related products. The researcher faces the selection problem at each of these phases.

Ideally, selecting video events from a larger source of information will proceed strategically at each phase. We say "ideally" because researchers do not always carry out selection or have control over selection at each research stage. For example, in his study of the Rodney King incident and trial, Goodwin (1994) analyzed *available* video of the incident. Although good research can be carried out with repurposed video that was collected without consideration of research goals, the researcher must consider how selection at previous stages affects analytical possibilities when video is used. For example, Leonard (2006) analyzed students' thinking and learning in video recordings of science classes gathered for the primary purpose of creating professional development materials. Presentation of her analysis included discussion of the strengths and limitations of her data sources, including conclusions she could and could not draw from them as a result of the way they were collected.

In cases in which the researcher does plan and carry out the collection of the video that is later analyzed, the selection issue in earlier stages of work is deeply intertwined with practical and technical skills, as well as ethical issues, related to planning and executing a video "shoot" in the field. These practical skills, which we suspect are not often taught in graduate-level courses, include (a) knowing how to choose and place cameras and microphones, (b) deciding when to start and end shooting, (c) deciding whether to shoot mainly wide angle or close up, and (d) making panning and zooming decisions in what is called *camera editing*. A discussion of these technical issues is omitted here, but Appendix A supplies practical advice, much of it directly related to the selection problem, for researchers using video recording to collect dense, close-to-reality information about ongoing human activity in complex learning environments. Ethical issues are addressed in a later section.

Here we propose additional ideas about how selection influences all stages of the video research process. To organize our discussion, we contrast two different purposes for making video selections: (a) to obtain sources for data extraction and analysis and (b) to create a narrative account of some phenomenon. In regard to both we distinguish between inductive and deductive approaches. A complexity we acknowledge and discuss is that good video research often blends these purposes.

Video as Sources for Data Analysis

The vast majority of educational research using video has involved detailed analysis of selected clips. Researchers selecting video clips for this purpose are often concerned with closely describing and accounting for the relative frequency of a type of event. They may need to determine an event's typicality or atypicality and the distribution of its occurrence: For example, does event X happen only in the classroom or also at home and in after-school settings? Within classroom type A (e.g., a math lesson), where and how often does event X happen? Does event X happen in every recorded instance of a small group in classroom A or only in some instances? And so forth.

We distinguish between inductive and deductive approaches to selection for data analysis. *Inductive approaches* apply when a minimally edited video corpus is collected and/or investigated with broad questions in mind but without a strong orienting theory. One generally begins by considering the corpus (or as much of it as possible) in its entirety, then considering it in progressively greater depth. The whole-to-part inductive procedure described by Erickson (2006) recommends repeated viewings of the corpus of interest in which multiple viewers reach agreement on major events, transitions, and themes. Abstract "intermediate representations" that describe the corpus as a hierarchical macroevent may be developed, as is discussed in the next section on video analysis. Such intermediate representations can help researchers strategically select events for deeper analyses that adequately cover major themes and include key participants and hence constitute a kind of representative sample from the macroevent. The conversations and nonverbal behaviors from such a sample, which are often transcribed, become the selection for deeper analysis.

A famous example of inductive contrastive analysis in the educational literature is Hugh Mehan's (1979) analysis of question–answer sequences in whole-group lessons in inner city elementary school classrooms. Mehan videotaped lessons in a range of subject areas during an entire school year, transcribed every lesson, and then identified every question–answer pair and every topically tied sequence of such pairs. He then showed that the vast majority of the instances were of the three-part sequential I-R-E form: **I**nitiation by the teacher (by asking a question the teacher knows the answer to), **R**esponse by the student (saying something related to the question), and **E**valuation by the teacher (concerning the correctness or appropriateness of the student response). What Mehan did that was methodologically rare at the time was to perform an exhaustive, contrastive analysis (Glaser & Strauss, 1967) on an entire corpus of instances. Every instance of a question–answer sequence in his data set was selected and analyzed.

A *deductive approach* is required when the researcher has a strong theory and clear research questions. Deductive approaches involve identifying or creating a suitable video corpus and systematically sampling from it to examine specific research questions. For example, Alibali and Nathan (2007) examined how a teacher used gestures in explaining mathematics to middle school students under the hypothesis that gesture is more prominent and important in introducing new topics. They selected video clips in which the teacher either reviewed old topics or presented new ones. Next they discovered facts about the clips related to the research questions. For example, how many gestures and what types of gestures did the teacher use? Finally, they developed a coding system to categorize the facts and calculate the frequencies of the occurrences, statistically comparing, in this example, "new topic" and "old topic" samples. This coding approach was then used to investigate future samples.

These examples illustrate cases in which researchers strategically selected video segments from an available corpus and used them for a specific analytic purpose. Such purposes differ in accordance with the theoretical orientation and research questions of the researcher. One researcher may select and examine clips to study mutually influential relationships between the gaze of the speaker and nodding reactions by the listener, as did Erickson and Shultz (1982) in *The Counselor as Gatekeeper*. Another person could be studying the same video to identify *implicit discourses*—value assumptions and student identity definitions—as part of the so-called hidden curriculum, as Jay Lemke (2001) did in his study of science classrooms. Or the researcher might select clips that provide data on students' transcription practices (Lehrer & Schauble, 2004), or teachers' representation and modeling practices in design-based science (Leonard & Derry, 2006), or scientific terms in the individual speech of students in small groups (Lynch, Kuipers, & Pyke, 2005), or critical conceptual leaps in children's mathematical growth

(Powell et al., 2003). In all of these examples, video clips were selected strategically and then further mined by researchers who used various analytical, coding, and/or transcription procedures to help identify and examine patterns in and across video data sources.

Selecting Video for Narrative Power

A quotation by Bruner (1986) helps elucidate the logic of selection for researchers who collect or use video primarily to tell an insightful story, such as a case study:

> Understanding is the outcome of organizing and contextualizing essentially contestable, incompletely verifiable propositions in a disciplined way. One of our principal means for doing so is through narrative: by telling a story of what something is "about." (p. 90)

Researchers taking a narrative approach may also work more or less deductively or inductively. Those with a strong theory or set of research questions may delineate the terrain in advance by designing interviews or protocols before stepping into the research setting. In the absence of strong theory, researchers may enter the research setting with broad thematic questions and work inductively. Inductively oriented researchers in particular may "hang out" at a research site to establish rapport with the participants before the video cameras are even used, seeking to become part of the culture without creating much distraction (Becker, 1998).

Narrative approaches derived from ethnography involve what Mary Catherine Bateson (1989) called *disciplined subjectivity*. The researcher makes his or her perspectives as transparent as possible. There is no pretense of objectivity because one main tool the researcher uses is his or her ability to learn and represent the culture. To build stories, a narrative researcher may select interesting and illustrative video segments soon after they are gathered throughout the study, at regular intervals, or toward the end of the study. A continual and interactive process of building the researcher's understanding, similar to how a film director watches rushes and makes selections, influences these selections. Current technologies afford repeated viewings and frequent returning to the data set to make more precise "cuts" of the video stream, increasing the possibility that the study changes during the process of selection. These activities increasingly sharpen the "eye" of the researcher. As selection proceeds throughout the project, the researcher is making decisions about why to choose one segment and not the one before or after, in accordance with a narrative structure that may be emerging.

This narrative approach differs philosophically from ones in which researchers strive, using minimally edited video footage, to systematically sample events as representative data sources. Many researchers who use the narrative genre select themes and then layer them in ways that are "thick," rich with description (Geertz, 1973). Their aim is not to make the complex simple, reducing data to theorem. The aim is to make the complex understandable. This is accomplished though a process of selectively organizing a research presentation (a story) into "digestible" chunks and then contextualizing them within a narrative thread that not only makes consumers of the research (viewers as well as readers) feel they were present but may also include participants as partners in telling the stories.

In selecting video chunks to tell a story or a piece of a story or to show what an instance of a given event phenomenon looks and sounds like, the researcher may look across a vast array of video records to find the most representative instance or instances—perhaps the most salient video chunks that best illustrate and represent one day or month of data collection. These mini-cases may become part of a larger movie or digital archive. Narrative-oriented researchers tend to underscore the importance of aesthetically pleasing images (Tobin & Hsueh, 2007) and good sound capture, taking advantage of the power of video to be a compelling narrative medium. And although a video chunk chosen for the purposes of representation or storytelling may not be selected primarily for its adequacy to support analyses that will provide evidence about the relative frequency of a particular type of interactional phenomenon (e.g., a kind of utterance, gesture, or math reasoning), the possibility of applying thick description to interpret a selected clip certainly exists. Even if selected because of its qualities as a narrative example, the clip might help define an interactional phenomenon under study and may arguably be a representative case of it.

To help distinguish the essential differences between selection for data analysis versus selection to support narrative, one can ask whether one's goal is (a) systematically selecting representative clips to help identify and document some naturally occurring pattern; versus (b) selecting events that best show or illustrate key events in a researcher's evolving interpretive narrative. Another rule of thumb is to determine the extent to which dramatic or aesthetic criteria are being considered. In choosing a clip for its narrative potential, what the clip looks and sounds like and its overall appearance according to aesthetic criteria may be of foundational interest and may help govern the selection. In contrast, when choosing a clip for its potential as a source of data that will be mined using coding and analysis procedures, what the clip looks and sounds like is not a major determining factor for selection; such considerations might even bias the research. This does not mean that data-mining researchers should ignore technical considerations that determine the quality of their sound and picture. But picture and sound that are merely visible and audible may be "good enough," aesthetically speaking, for data mining.

Projects in which narrative and illustrative concerns are at the fore are exemplified by some teacher professional development Web sites such as Teachscape (2009), in which developers make available exemplary or illustrative video cases of teaching practice in the disciplines at different grade levels. Other examples from teacher professional development are the case studies of student thinking selected for *Seeing the Science in Children's Thinking* (Hammer & van Zee, 2006). The video collection located on a Web site cross-referenced to the book *Points of Viewing Children's Thinking* (Goldman-Segall, 1998) represents one of the earliest examples of digital video used in this manner. The Carnegie Foundation's Gallery of Teaching and Learning (Carnegie Foundation for Teaching and Learning, n.d.), on which teachers posted thoughtfully developed cases of their teaching scholarship, illustrates selection of video to support personal narrative.

Blended Research and Selection

Although distinguishing two purposes of video selection is conceptually useful for helping researchers reflect on the meaning of their decisions, it is important to realize that these purposes can support each other. For example, in his book *Talk and Social Theory*, Erickson (2004) used two purposes of selection simultaneously. He presented four video examples on a Web site—one for each of four chapters in his book. These clips served as illustration and narration. But they were also treated analytically (and microanalytically) in the book chapters, with transcription of verbal and nonverbal behavior presented in differing grain sizes and supporting discussions of issues at differing timescales.

The Trends in International Mathematics and Science Study (TIMSS) video project (Stigler & Hiebert, 1999) represents an additional way in which the two categories of selection overlap conceptually. The huge TIMSS video corpus has been selected from for professional development purposes and to support cross-cultural studies of teaching. But *sampling representativeness* has been an overriding concern in selecting for both purposes. In the TIMSS project, a formal sampling logic governed the selection of video in stages of research that occurred before recording (planning what countries and schools to sample). During recording, rules for sampling behaviors guided the video selection that occurred at camera editing time: The camera operator was instructed to adopt the perspective of a "good student" paying close attention to what was happening in class.

Even in situations in which developing a narrative is the main goal of selection and aesthetics are important, issues of selecting *fairly* to represent culturally prototypical cases of practice have come to the fore. A good illustration of this issue was discussed by Miller and Zhou (2007), who described a decision to stop showing aesthetically and conceptually pleasing clips of mathematics teaching in China when it was discovered that the methods illustrated were considered outdated and atypical of current teaching. A similar illustration is found in Tobin, Wu, and Davidson (1989), who videotaped preschool classrooms in three countries and produced edited videos depicting a "typical day" in each setting. They then showed them to teachers and principals in each country for confirmation that the events shown were typical and representative. Where there were disagreements about what was representative, these were described in the video cases.

These examples help lead to an important point: When researchers collect video that will become a public corpus or that will later be shared and selected from by other researchers or even used again later by the original researchers, the original researchers will not, at the time of collection, have anticipated all potential and future uses of the video. To help ensure that future users of the video will not inadvertently commit errors of selection bias—for example, by treating a video corpus or clip as though it is a representative sample of practice when it is not—it is necessary to document as fully as possible archived video images with information about when, where, and how they were collected. Such documentation should permanently live as metadata with the video corpus and be available to and used by researchers who later select from it. Such context documentation is also best practice in a broad range of other field sciences, including ethology, field ecology, and geology. However, documenting a video corpus raises numerous issues, as described later in our discussion of ethical concerns related to video data collection.

Another type of blending narrative with data-mining purposes occurs when video clips that were originally selected to support narrative storytelling and that may have been nonrepresentative samples are subjected to deeper analysis. Here we refer to scientific analyses of video performed by researchers, distinguishing these from less formal analyses of the type conducted by teachers who participate in lesson study or video clubs as an approach to professional development (e.g., C. Lewis, 2002; Sherin, 2004, 2007). The practice of analyzing video clips without making strong claims about what categories of phenomena these clips represent more broadly is supported by an alternative viewpoint toward selection held by many conversation analysts (Atkinson & Heritage, 1984; Schegloff, 1997); at this point in the study of human social interaction, researchers know so little that they should, as a matter of principle, presume that each thing that happens interactionally is of equal potential interest and importance. This approach places few constraints on selection: Almost any minimally camera-edited video clip of human social interaction is worthy of being investigated. This method takes a very different perspective from that held by researchers whose selection heuristics are guided explicitly or implicitly by probabilistic concepts of frequency and representativeness and also from that held by researchers whose selections are based on their significance and meaning within a narrative account.

Selection Strategies: Concluding Comment

Selective emphasis is a fundamental and unavoidable process that strongly shapes video research at every step during all phases of inquiry. Here we attempted to achieve a better understanding of video research by uncovering and reflecting on the problem of selection. To help frame the problem, we distinguished between two ways of thinking about selection that reflect different ways in which researchers use video records: (a) to locate and analyze data for the purpose of finding patterns within and across events; or (b) to use video clips more holistically to support an evolving narrative. In practice, many research projects blend both selection logics. Our goal here was to heighten awareness of both types of selection as well as their complementary natures. Dewey (1958, as cited in Clancey, in press) argued that selective emphasis is inevitable whenever reflection occurs. Such selective emphasis is a negative force only if, in the researcher's thinking and actions, the presence and operation of choice are concealed, denied, or disregarded.

ANALYZING DATA DERIVED FROM VIDEO RECORDS

We now strive to create a more coherent conceptualization of how to conduct an in-depth analysis when selected video records are the primary data source. We address related issues in the stages of planning, theory inquiry cycles, viewing practices, intermediate representations of video records, and ways in which researchers use video records to create data and make theoretical or empirical claims. How one approaches an analysis depends crucially on one's theoretical commitments; the specific research questions one is pursuing; the rhetorical situation for which the video analysis is being conducted; and the practical constraints of time, money, and personnel. In video analyses of high quality, the researcher makes a convincing case that his or her analytic choices were sufficiently responsive to these considerations. Reliability and validity issues of all kinds (internal, convergent, external, descriptive, interpretive, and theoretical) apply to video-based data just as they apply for any other kind of quantitative or qualitative data analysis. Potential criticisms from the research community about the generalizability of findings from video research can be countered by paying explicit attention to the logic of one's inquiry, including one's approach to selecting or collecting records, and by articulating the processes used to create explanations and generate claims. What is clear is that performing analyses with video is an iterative process that involves moving back and forth among the process of video selection; one's evolving interpretations and hypotheses; and a variety of intermediate representations for discovering, evaluating, and representing the video data for oneself and others.

Advice Based on the Wisdom of Practice

The experiences of researchers using video as a tool for inquiry have yielded a number of models and concrete suggestions for conducting video analyses that are particularly useful for beginning researchers (e.g., Barron, 2006; Cobb & Whitenack, 1996; Erickson, 2006; Powell et al., 2003). We begin by summarizing some of these insights and pointing readers to sources where they can learn more.

Start with guiding questions. Ideally, the collection of research video is guided by a plan and a set of research questions that are based on the researcher's familiarity with the phenomena being studied, although situations also arise in which video that has already been collected and archived is analyzed (e.g., Goodwin, 1994; Leonard & Derry, 2006). In either case, planning is particularly helpful when the researcher is new to video analysis. The amount of detail that can be captured in video recordings makes them a powerful resource compared to what the human observer can record in real time, but it also makes them challenging to work with. Turning raw video records into useful data is enormously time consuming. Thus, it makes sense to go into a project with theoretically motivated questions that originate from the research literature or observations. Good orienting questions help the researcher maintain a perspective that prevents him or her from getting lost in the details that video records include.

Reflecting on what theoretically motivated questions might be pursued may fundamentally influence strategies for data collection. For example, many investigators have found it fruitful to combine video records with other forms of data, including data from performance assessments, interviews, and surveys—or, in other kinds of studies, psychophysiological data (heart rate, galvanic skin response) and eye-tracking data. Field notes, photographs of the surrounding field of action, copies of documents involved in participant interactions, and products created by groups can enrich the data derived from video records and offer opportunities for triangulation across sources of evidence. The process of developing questions can help the researcher decide what needs to be captured on video and with other methods. For example, if the question of interest is the role of material artifacts in the alignment of the attention of group members, the plan might include collecting the artifacts (such as diagrams created by collaborators) and having the camera positioned to film them in use. In the increasingly frequent situations of interactions mediated by computer applications displayed on screens between participants, synchronized screen recording in addition to video of the participants interacting is advisable (e.g., Stevens, Satwicz, & McCarthy, 2008). Appendix A provides practical and technical advice related to some of these issues.

Expect unanticipated phenomena. At the same time one is working with guiding questions, it is important to also remain open to discovering new phenomena. For example, Engle, Conant, and Greeno (2007) were interested in conceptual change, so they designed a data collection plan that included pre- and post-assessments intended to measure changes in students' conceptual understanding. They also collected video data of classroom conversations that were likely to have generated conceptual growth. But during analysis, some totally unanticipated findings emerged. The researchers addressed the good questions they started with, but the novel phenomena were theoretically fruitful. Formulating and answering questions does not preclude additional discovery-oriented work with video records. In fact, this is one of the valuable properties of video records—they can be revisited for continued learning and analysis at different times, with different viewpoints, and by different researchers. This is one of the primary arguments for moving toward more open data access than is typical today.

Develop social practices for viewing. One advantage of video recordings as a source of data is that they can be viewed multiple times in different ways, with different people, at different times in the history of a research project, and even by different research groups. Investigators can strengthen their research findings by coordinating what they learn from multiple viewing opportunities. In the early stages of a video analysis, before interpretations of events become fixed in the researcher's mind, it can be quite helpful to share a key video segment with a group of other researchers to gather multiple interpretations of the events, to surface salient dimensions for analysis (which will often vary by researcher disciplinary background), and to brainstorm potential issues to investigate further (Jordan & Henderson, 1995). The video segment can be viewed and re-viewed by the group to look for data consistent or inconsistent with initial hunches. Watching the videotape at a speed that is slower or faster than normal, only listening to the audio, or watching the video without audio can also help focus viewers' attention on particular aspects of interest (Erickson, 1982). Group viewing can be used in later stages of work to determine whether multiple researchers notice similar phenomena (e.g., Engle et al., 2007). Finally, it can sometimes be helpful to have people who were recorded watch the video in the presence of the researcher to provide their interpretations of what was going on. It is preferable to obtain participant involvement as soon as possible after recording and without asking leading questions (Ericsson & Simon, 1980; Jordan & Henderson, 1995).

Practices for Data Creation and Analysis

Video provides rich records of interactional phenomena, including eye gaze, gesture, body posture and proximity, content of talk, tone of voice, facial expressions, and use of physical artifacts, as well as between-person processes such as the alignment and maintenance of joint attention (Barron, 2003). Because this complexity makes it easy to become lost in detail, explicit strategies for focusing the attention of the analysts are needed. Strategies are also needed for establishing the content of the tapes and making decisions about how to represent the phenomena included within them.

Erickson (2006) provided three sets of guidelines, each reflecting fundamentally different approaches to inquiry. Briefly, he described (a) a whole-to-part inductive approach, in which social viewing and re-reviewing are used to identify patterns in data for which there are no strong orienting hypotheses, predictions, or theories; (b) a part-to-whole deductive approach that involves looking for specific types of events and that is appropriate when research is driven by strong questions, hypotheses, or theories about those events; and (c) the manifest content approach, in which interaction focusing on particular pedagogical or subject content is selected out and examined. Such suggestions are very helpful for the beginning or experienced researcher. There are numerous other ways in which to analyze and develop an understanding of interactions captured in video, and we now provide additional ideas to draw upon.

Representations for Data Selection and Pattern Finding

Creating intermediate representations of video records is important because such representations allow the researcher to identify which segments to analyze and to begin to see patterns within and across segments. Making transcripts of talk and nonverbal information is common. However, experienced researchers often take other preliminary steps to understand their data set and to initiate the process of pattern finding. Developing preliminary representations can help the researcher decide what should be transcribed and at what level of detail. Here we describe several approaches to developing representations of video data, including transcripts, and the variety of decisions that are involved with each approach.

Indexing. As emphasized in Appendix A, the first possible time to view video is while it is being collected. If a researcher can be present during recording, then he or she can make time-indexed field notes that provide a basic outline of the events or possible examples of phenomena of interest that occur while also potentially filling in relevant complementary information that is difficult to discern from the video itself. If the researcher cannot be present during recording, it is very helpful if he or she can watch the video soon after it was recorded in order to create a *content log*, which, like field notes, will provide a time-indexed outline of the events on the video. Content logs can be extremely detailed, consisting of a description of major events that took place for each brief standard unit of time (e.g., 3 min), or they can consist of a several-sentence description of the content of a whole hour of instruction. Field notes and content logs allow the research team to develop a quick sense of the corpus of data and to facilitate the selection of episodes for further detailed analysis. This kind of indexing should be distinguished from systematic coding. *Systematic coding*, as we discuss later, is best done after extensive work is completed to establish the meaning of codes and the central units, such as events or turns in a conversation, that should be coded.

Macrolevel coding. Because transcription is costly and time consuming and not always suited to pattern finding, video researchers often invent other ways to summarize records. For example, Ash (2007), who studies family conversations in museums, begins with a representation she calls the *flow chart*, which catalogues a family's museum visit from start to finish, including any pre- or post-visit interviews. The goal is to mark major events and the occurrence of conversations about biological themes. Topics and themes can be coded from this representation to compare different visits made by the same family or visits made by different families. For Ash, the flow chart representation is also key for selecting the data for her second macrolevel unit of analysis—the significant event. Significant events are selected based on four criteria: (a) They have recognizable beginnings and endings (usually they take place at one exhibit); (b) they have sustained conversational segments; (c) they integrate different sources of knowledge; and (d) they involve inquiry strategies, such as questioning, inferring, and predicting. Ash's final level of analysis involves microlevel examination of the interactions occurring within selected significant events. For example, Ash and her team use discourse analytic frameworks to study how an idea develops over time.

Narrative summaries. Other researchers use narrative accounts that may be informed by or contribute to additional forms of analysis (e.g., Cobb & Whitenack, 1996; Powell et al., 2003). Angelillo, Rogoff, and Chavajay (2007) conducted a video study that compared mother–child interactions in four distinct cultural communities. Their first step was to generate descriptive, narrative accounts of each 1.5-hr videotaped home visit during which mothers helped their toddlers learn about novel objects. These were not event logs but instead were lengthy (as long as 30 pages) descriptions of events. These descriptive accounts were used to help the rest of the research team visualize the sequence of interactions and to capture the purposes and functions of action and dialogue.

Diagrams. Other researchers summarize aspects of video records using still frames or diagrams. For example, in a study that investigated patterns of activity between Guatemalan Mayan mothers and children completing puzzles, the researchers' goal was to categorize patterns of joint attention, mutual orientation, and ways of distributing work (Angelillo et al., 2007). The team created a diagramming method that allowed the researchers to characterize types of coordination around shared tasks that involved multiple people. The diagrams were then used to help code 1-min intervals of video.

Transcription. Although there are exceptions (e.g., see Angelillo et al., 2007), during the process of video analysis most researchers produce transcripts that represent some portion of the events recorded in their video. Initial transcripts may help researchers flesh out from their field notes or content logs what occurred in a particular segment of video in order to decide whether and how to pursue an analysis (Jordan & Henderson, 1995). In later stages of research, transcripts are iteratively revised while video analysis proceeds, until the transcripts eventually provide a reliable record of what the researchers view as the most relevant aspects of the video for their research questions (e.g., Engle et al., 2007; Mischler, 1991). Through this process, transcripts become key data that can be used directly for additional coding, interpretation, or creation of other analytical representations. When research is reported, transcripts must be edited for public consumption (e.g., Du Bois, Schuetze-Coburn, Cumming, & Paolino, 1993), although transcripts developed during analysis often include standardized symbols and spatial conventions to represent various aspects of communication and action such as gestures, intonation, pauses, and speaker overlaps. Whether explicitly intended or not, the transcript convention used embodies theoretical value judgments about which recorded events are important (Lapadat & Lindsay, 1999; Ochs, 1979). There is no such thing as a "complete" transcript that captures the full complexity of all verbal and nonverbal events. Consider the relation of the transcript to actual events by analogy of the relation of a map to the world it represents. There is no sensible "complete map." Borges's (1999) story "On Exactitude in Science" parodies this concept when he imagines an empire where only a map on the same scale as that empire will suffice for precision.

There are many existing—and in many cases competing—conventions for how one might transcribe different aspects of the social interactions captured on video. In Appendix A of the report *Guidelines for Conducting Video Research in Education* (Derry, 2007a), we compiled a list of common transcription choices, summarizing their features and their strengths and weaknesses. This appendix also contains references to work containing examples of these transcription conventions. Typically, researchers adapt existing conventions in ways that make sense given their research questions; their theoretical commitments; and practical constraints such as available time and personnel, the audiences for their work, and the systematic availability and accessibility of information in the video record and other data sources. The important thing is to explain how one's own decisions about which transcription conventions to use make sense given these various considerations.

Coding, Counting, and Quantitative Analysis

Because the method of analysis and how a video analysis is reported are related, in this section we combine a discussion of analysis and reporting, referring to examples in the published literature. Methods of analysis in which video records are coded are rooted in practices of *disciplined observation*, a core feature of scientific methodology. Independent of the advent of video technologies, social scientists developed approaches that allowed them to document, analyze, and report human behavior observed in natural contexts. Systematic observational approaches relied on preestablished coding schemes and were designed to yield reliable judgments by independent observers. For example, early studies of children's play often relied on repeated short samples (Goodenough, 1928) in which a child would be observed for 1 min a day and the researcher would code the child's activities into one of six mutually exclusive categories (Parten, 1932). The development of statistical approaches for determining interrater reliability was a key innovation. Before video, these methods required that the focus of inquiry and coding systems be fully worked out before data were collected and be simple enough for two or more observers to achieve interrater reliability. Video relieves these constraints, allowing analysts to develop complex coding systems over time.

Video researchers often develop systems of analysis over the course of multiple research projects (e.g., Powell et al., 2003). Ash (2007) articulated the changes that have occurred in her coding system and described the system that evolved, which she calls Tools for Observing Biological Talk Over Time. Through careful analysis of the talk of families and consultation with biologists, psychologists, and educators, Ash and her team have developed a system that can be used across projects and by other teams. They have gone through more than a dozen iterations to reach what they consider to be a stable and generative analytical system.

Like the processes of generating research questions or creating intermediate systems of representation, the development of a coding approach benefits from iterative cycles of work, distributed expertise, and moving across different levels of analysis. For example, Angelillo et al. (2007) described an approach to investigating patterns of shared engagement that combines qualitative and quantitative methods. The core of the process involves close ethnographic analysis of a few cases in order to build up a coding scheme that is based on the observed phenomena and that can then be applied to multiple cases. They illustrated this approach in their study focused on cultural variations in mothers' and toddlers' contributions to understanding novel objects across four culturally distinct communities. The research team approached their analyses with some tentative hypotheses about the kinds of interactions that might differ across the four cultural groups, for example the relative reliance on words versus nonverbal demonstration. However, as is the case with many video studies, the interactions caught on video led to the discovery of new phenomena, such as differences in ways in which the mothers from different cultures motivated engagement. Once these phenomena were identified, the team worked to refine the categories so that the new instances of the phenomena could be reliably coded.

Studies of interactions between mothers and children or between intimate partners provide exemplars of video research using coding and quantification. Video studies have made important contributions to researchers' understanding of early emergence of sophisticated social awareness in infancy and the bidirectional influences between caregiver and child (M. Lewis & Rosenblum, 1974; Stern, 1977; Trevarthen & Aitken, 2001). In the 1970s, Gottman's influential work developed methods that involved videotaping interactions of couples at a specially constructed "talk table" while the couples also rated aspects of their own communication. This research was aided by other influential applications of video, namely the study of facial expressions of emotion, which led to a coding system called the Facial Action Coding System (Ekman, Friesen, & Ellsworth, 1978). More recently, video research on couples' interactions has moved out of the lab and into the home to provide accounts that are more ecologically valid (Gottman & Notarius, 2002).

Erickson (1977, 1982, 1986) has written extensively about possible roles of quantification in qualitative research and has a useful discussion of the synergies between approaches. He has emphasized that determining what to count is more challenging than doing the actual counting. Other excellent discussions of the development and use of observational coding schemes and associated statistical techniques include a primer on the topic of sequential analysis by Bakeman and Gottman (1997) and an article by Chi (1997) that focuses on the coding of verbal data.

Despite the number of studies that use coding approaches, it is by no means universal that data derived from video records are coded in a way that can yield quantitative data. Many researchers prefer to focus on the development of rich examples and therefore do not count types of events within or across cases. One common way of reporting a video analysis in a publication is to provide a "play-by-play" description in which interpretations of episodes that follow one another in time are presented sequentially. Play-by-play analyses are particularly effective at showing how the sequentially developing context relates to what happens next. When supported by rich transcripts, these kinds of analyses are also particularly good at demonstrating how multiple actions and people collectively produce phenomena. In one extension of play-by-play analyses, a researcher might analyze selected episodes that all focus on a particular topic or other issue over the course of days, weeks, or even months to show how that issue was transformed over time. Examples of this approach in the published literature include Ochs and Taylor (1996); Koschmann, Glenn, and Conlee (1999); and Wortham (2004).

Progressive Refinement of Hypotheses

A recent volume on video analysis in the learning sciences contains chapters by a number of research groups that include details on video practices (Goldman, Pea, Barron, & Derry, 2007). In this volume are numerous examples of studies that interweave top-down planned analyses and the reporting of unanticipated phenomena. Some authors describe processes that resemble an approach to qualitative research more generally called analytic induction (Znaniecki, 1934). In *analytic induction*, a few cases are explored in depth and explanations are developed. New cases are examined for their consistency with the explanations, and when inconsistencies are found, the explanation is revised. A methodological paper by Cobb and Whitenack (1996) discussed a similar logic. In line with Glaser and Strauss's (1967) approach, their process for analyzing video of children's mathematics development in small-group settings over time involved constantly reconciling provisional analytic categories with subsequent data and newly formulated categories. They described their work as a "zigzag between conjectures and refutations" (p. 224).

A related approach offered by Engle et al. (2007) is *progressive refinement of hypotheses*. In this approach a general question is framed, and video records are collected in an appropriate setting. Once records are collected, more specific hypotheses are formed after some viewing of the records. These hypotheses are then examined in relation to other aspects of the data set, and more complete explanatory hypotheses are developed. Engle et al. argued that multiple iterations through hypothesis generation and evaluation lead to greater robustness of findings and increased likelihood that findings might be replicated in other contexts.

Reporting Results

Although some researchers have created special multimedia journal issues that include some video as part of the publication (e.g., Koschmann, 1999; Sfard & McClain, 2002), in most cases the video records will be left behind in the reporting phase of the project. What was observed must be re-represented in static text or graphics for journal publication. Coding and subsequent quantification is a common approach to reporting results. However, although the ability to code behaviors can rest on the well-developed techniques and methods described earlier, quantification does not allow the researcher to communicate how an interaction unfolds across time in all of its complexity. Narrative description is a method of representation that better describes the dynamics of interaction; however, narrative accounts are often less credible to many experimentally minded social scientists. One solution to broadening acceptance of a video analysis is to use more than one method of representation when reporting the research. For example, Barron (2000, 2003) used quantitative methods to find response patterns that reliably differentiated among more and less successful collaborative groups. However, the ways in which these sequences unfolded for individual groups differed in some important ways that were masked by the quantification. Thus, Barron combined what Bruner (1986) described as a paradigmatic approach (coding and statistical analysis) with a narrative approach (which preserved the sequence of interactions). Barron's narrative approach used three types of representations to convey the complexity of interaction: *transcripts* to illustrate key aspects of dialogue; *behavioral descriptions* that conveyed aspects of the interaction such as facial expression, tone, and gesture; and *still frames* to further illustrate the body positioning of interacting students at key points.

Erickson (2006) provided a particularly strong argument that readers of analyses should come away not only "tree-wise" but also "forest-wise" (p. 185). That is, it is not enough to provide rich examples; analysts must also provide a sense of the broader sample and of how typical or atypical the instances presented are relative to some larger corpus of data. Our discussion has suggested ways of communicating levels of analysis as well as interrelationships among the levels. The problems of re-representing the complexities in video are not trivial, and the video research community is in the beginning stages of figuring out as a field creative ways to achieve compelling representations of complexity.

Analyzing Data From Video Records: Concluding Comment

In summary, video analysis can range from discovery-oriented approaches with the hope of revealing unanticipated phenomena to top-down approaches that use records to identify and code events that have been mostly conceptualized prior to data collection. Researchers starting to plan a project that will use video records should focus first on theory-driven questions and develop concrete plans for a first pass through the video records. Having good questions will help maintain perspective and prevent the researcher from getting lost in detail. At the same time, the researcher should anticipate new discoveries and be ready to articulate questions that can be followed and refined and tested during multiple passes through the video records. These passes can be made most fruitful by using intermediate representations. Researchers should expect to have to engage in multiple cycles of analysis. An explicit multistage analytic approach can strengthen the likelihood of generating strong findings that are both reliable and valid.

TOWARD SOCIAL TECHNOLOGIES FOR VIDEO RESEARCH: AN EMERGING AGENDA

Because video research is organizationally and analytically complex, researchers benefit greatly when common boundary objects include good technological tools to support their work. However, selecting good tools for video research is not easy. There are many choices, and many new technologies are still developing, including several cyber infrastructure projects to watch and possibly become involved in.

Here we address the role and importance of present and future technologies in the sharing and reporting of video research. These include technologies for video analysis, technologies for supporting video case development and sharing, models for sharing video in reports of research, metadata schemas, collaboratories and virtual repositories, and practices addressing legal and ethical issues related to video sharing.

Before discussing these technologies it is important to emphasize that video research workflow is a highly iterative and often social process, not a straightforward linear progression from capture to transcribing to coding to reporting. Even as video records are collected in the field, the video researcher begins the process by pulling video recordings into some kind of order, from the simple act of labeling them for easy retrieval later to the much more intricate activities that add the value of interpretation to video records. The researcher may chunk the video records into segments that are defined by event boundaries, time markers, or a variety of semiotic considerations, as previously discussed. And the researcher continues by marking video segments of interest, creating transcripts at different levels of detail depending on the purpose at hand, and developing and using categories that he or she considers useful, in a recursive manner. These activities involve an ongoing accumulation of research identified through processes of finding, tracking, searching, and browsing, at best yielding a deepening analysis of the human activities that have been recorded. The researcher marks, transcribes, and categorizes a little; analyzes and reflects a little; searches and finds a little; and so on, in the recursive loops that define such knowledge-building activities (analogous to the writing process). In essence, there are close interdependencies between the activities of video record decomposition (e.g., segmenting, naming, coding) and recomposition (e.g., making narrative case reports, collecting instances of commonly categorized phenomena, making statistical comparisons of chunked episodes). Then the workflow moves on to presenting and sharing video analyses in a variety of formats. Such sharing may be formative, as one collaboratively develops and comments on a developing video analysis, or it may be a summative account as the video analysis is published (e.g., in a journal, on a DVD, or on the Web) and commented on by others in the community. To close the loop, the substantive insights from specific video research workflow activities might influence the next cycles of video research workflow in the field. This basic idea is depicted in Figure 1.

Most available video research technologies support only certain phases of this workflow process. Accumulating findings from a workshop that brought together leading video researchers in the learning sciences and teacher education, Pea and Hay (2003) identified 10 different functions of video research that are supported by different tools: (a) video acquisition, (b) chunking, (c) transcription, (d) wayfinding, (e) organization and asset management, (f) commentary, (g) coding and annotation, (h) reflection, (i) sharing and publication, and (j) presentation. However, most tools developed and used in video research groups have focused on developing only a few of these capabilities, and, as several examples illustrate, the tools vary considerably in how well they support these functions. For example, Video-Paper Builder is designed primarily to facilitate sharing and reporting through the creation of Web-based "video-papers," educational research publications that incorporate video clips (Beardsley, Cogan-Drew, & Olivero, 2007; Nemirovsky, Lara-Meloy, Earnest, & Ribeiro, 2001). The Computerized Language Analysis (CLAN) tools that MacWhinney (n.d.-a, n.d.-b) and colleagues have developed for TalkBank provide an exceptional suite of transcription, coding, and annotation tools but are not oriented to supporting reflection, sharing, or commentary. In contrast, the Teachscape platform for providing video case studies of exemplary teaching supports chunking (by highlighting sections of each of the

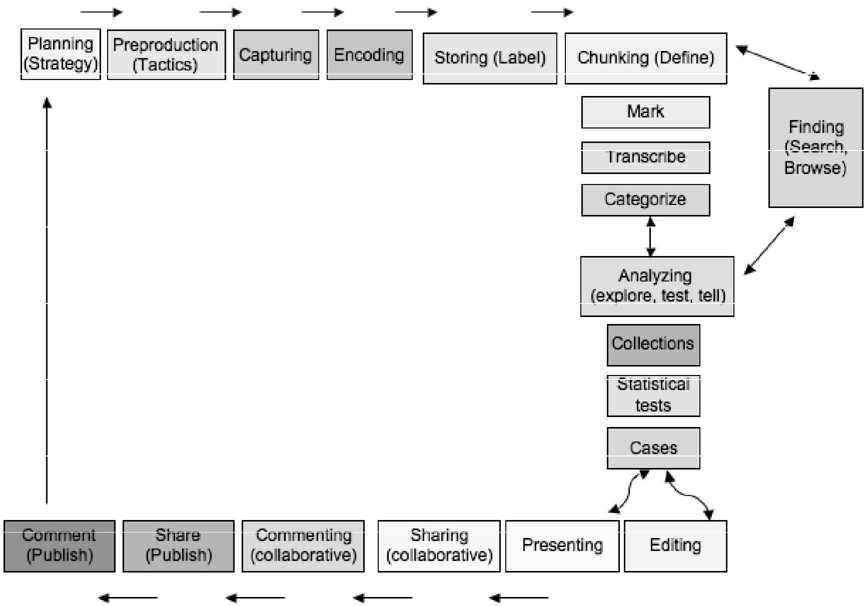

FIGURE 1 Diagram of video research workflow processes (Pea & Hoffert, 2007).

videos for particular instructional purposes) and reflection (by supporting study groups of teachers who use the online community features at home or in work groups at school; e.g., Lu & Rose, 2003). But the Teachscape platform does not support the development of coding schemes, nor does it provide transcription tools, because it was designed more as a teacher professional development environment than as a support for basic research.

Here we discuss key ideas and issues related to the development of analytic and social tools for video sharing and analysis, an important part of the current learning science research agenda.

Current Technological Tools for Video Research

The tools that researchers choose significantly shape their work. The tools researchers *choose to develop and make available* to other researchers may significantly shape the work of the larger learning science community.

There is a long history to video annotation and analysis systems, which have been under development for at least two decades (e.g., Harrison & Baecker, 1992; Mackay & Beaudouin-Lafon, 1998; Roschelle, Pea, & Trigg, 1990). Such tools help video researchers create and organize video collections, create transcripts, chunk and annotate video clips, search video banks, develop coding schemes, and create summary reports that support analysis. Video researchers must each investigate this issue and put thought into acquiring or developing a toolkit that meets the constraints of their budgets, can be accessed or supported from their sites, adequately supports their research practices and goals, and, ideally, helps involve them in a supportive community of users. In Appendix B we highlight a few possible existing choices of tools for video research. Many others exist, and development of new tools will continue apace as an important part of the emerging video research agenda.

Formats for Sharing Video Research

In addition to tools, the field needs models for different ways in which video data can be shared, commented on, and reported in concert with print and other media. Moving beyond the standard journal-reporting formats discussed in the previous section on video analysis, we provide here several exemplars of technical practice that include video data in the public reporting of research. Each one is regarded as a classic by many researchers, and each was produced by a senior researcher with substantial experience with the challenges of reporting with video data.

The first example is drawn from a "multiple analysis" project: the same video data independently analyzed by several leading researchers with differing theoretical perspectives. The analyses were published as separate articles in a single issue of the journal *Discourse Processes* (Koschmann, 1999). The issue contained an introduction by the researchers whose work generated the original video data and commentaries by additional researchers, much as in a conference symposium with discussants (which had in fact taken place before the writing of the final articles). In this case the data made available to the article authors—consisting of 6 min of a single-camera video record—were provided on a CD-ROM included with the distribution of the printed issue of the journal. The CD-ROM also included a detailed transcript and a drawing of the physical setting.

The second example is the distribution, again on CD-ROM but in a limited way and primarily to interested scholarly colleagues, of parallel print and multimedia versions of two research publications by Goodwin (2000, 2003). Provided were portable data format (.pdf) versions of the printed articles and more complete versions that included links to video clips, in-line image stills, links to audio-only files, and graphical markups of transcripts and stills showing relationships (e.g., with drawn arrows). These pioneering multimedia representations of interactional complexity (developed with Marjorie Goodwin) made the relationships between data and argumentation far clearer than was possible in the printed versions, which included only textual transcriptions. This model creatively used evocative still-frame capture from a video recording to represent vital aspects of an interaction sequence. Although Margaret Mead and Gregory Bateson pioneered the use of photographic stills from film for ethnographic analyses in the 1930s and 1940s, the

Goodwins' techniques illustrate many unique ideas for marking out and labeling the central properties of images used in analyses. A sense of the hypermedia style of the multimedia versions can be gained from browsing Goodwin's (n.d.) Web site.

The third example is the availability on the Web site "Points of Viewing" (Goldman-Segall, n.d.) of brief video clips keyed to pages in the text of *Points of Viewing Children's Thinking* (Goldman-Segall, 1998). At the bottom of relevant pages in the printed book are very small video stills representing the action analyzed in the text. Relevant video clips can be accessed on the Web site, and site visitors may add commentary. The commenting facility exemplifies the epistemological stance of the research: The meaning of what one sees depends on the perspectives (spatial and cultural) from which one views it.

In each of these cases an important criterion is met: *Make available to the audience of the research report a sufficient sample of the video data on which the report's argument is based to allow the audience to assess the quality of the argument based on the data.* Experienced researchers agree that just as the map is not the territory, so the transcript is not the video (nor the video, the event!). Even if research results depend often on analyses of transcripts, access to a sample of original video allows scholarly peers to assess the results of transcription and to place analyses in the wider context of features of the video-recorded event that may not have appeared relevant to the original researchers. Making available a sample of original video is a less stringent standard than making all data available for reanalysis. The purposes are to enable researchers to more clearly convey the evidentiary basis of their arguments and to permit a closer assessment of the work reported.

Given the limitations of print publication and of material distribution of media (e.g., through DVD), it seems likely that this guideline for good research practice—making available a sample or original video—can best be met in the near future by use of online multimedia supplements to versions of printed research reports. More broadly, researchers need to share and debate cases of research involving video analyses reported in standard journals, as well as models for alternative and experimental formats (e.g., good Web sites, such as those previously described, that incorporate video cases, analyses, and publications). It is heartening to see the "proof case" in the recent emergence of Web video journals to communicate the minutiae associated with specific scientific techniques such as those of cell biology (see *Journal of Visualized Experiments*, 2008). Another example is SciVee (2009), a Web 2.0 site that enables researchers to combine video with documentation and data in a media-rich format, making their research more accessible throughout the research cycle.

There are also vital roles to be played in the learning sciences by the early release of multimedia findings in a formative mode (Pea, 1999) before traditional print archival publications are developed in a summative mode, as in the justly famous case of e-prints in fields such as high-energy physics. The analogue in video research may well be draft releases of video analyses that are shared openly with a community (such as members of a collaboratory) before archival versions of these video analyses are published either in print, online, or in DVD media. It is now possible to upload and publish digital video recordings so that they can be played in streaming media or downloadable format not only from a Web site devoted to video (e.g., Google Video or YouTube) but also using a video player that can be easily inserted into blogs or social networking sites. Thus, the researcher can create a composite video analysis that references multiple clips that may reside on multiple servers and be provided to observers as, for example, a WebDIVER composite remix or "mashup" of the video clips (Pea, 2006). Such flexibly adaptive uses and reuses for video data and analyses raise important issues concerning attribution, standards for reuse and remixing, protocols for protecting human subjects and the like, which are discussed later.

Sharing Video as Data Sources for Research

As noted previously, it is good practice to make a sufficient sample of video available with published research based on those data, which enables readers to make certain kinds of judgments about the data and how they are being used. But it is also good practice to find ways to make a larger fraction, or even the complete corpus, of relevant video sources available to peer researchers for reanalysis. Reasons for doing so include making it possible to subject claims based on the data to scholarly debate and enabling other researchers to benefit from the time—and in many cases funding from the public's tax monies—invested in acquiring the data. In the past, making video data available was not practical because of the difficulty and expense of copying and distributing analog videotapes. Such drawbacks no longer apply to digital video data, and in the near future even very large video files will be made available online. Already, Amazon, iTunes, Netflix, and YouTube stream or download legally entire feature-length movies online. In this section we consider what boundary objects must be further developed to support such sharing for video research.

Metadata schemas. As interesting as the distinct video tools described previously are, perhaps the most important lesson derived from our workshop and conference on video research was that the usefulness of such tools is limited without effective metadata schemes (Pea & Hay, 2003). Unless metadata coding and associated XML[3] schemas are used for the purpose of exposing such coding work

[3]XML (Extensible Markup Language) is a general-purpose extensible language specification recommended by the World Wide Web Consortium as a free open standard for creating custom markup languages. XML schemas are designed for supporting information systems to share structured data over the Internet.

to browser search, analyses developed with any of these tools will be isolated in data islands that can only be used and understood within the tools and projects in which they are created. This is a serious problem.

Examples of the type of work on coding metadata that video researchers in the learning sciences must undertake are provided in Appendix B. Although the challenges affiliated with establishing broadly applicable metadata for video analyses and video cases in learning and teaching, including the need for generic and discipline-specific metadata categories, are significant, such efforts are needed to achieve the broader goals of establishing distributed research teams that can communicate about their video data productively.

Virtual repositories and collaboratories. Here we consider the prospect of sharing video through *digital video collaboratories* in which multiple researchers who are distributed across time, location, discipline, and hardware platforms can upload video files as common resources for examination. Although an individual researcher or group might have a sizeable collection of digital video assets, repositories associated with video collaboratories are envisioned on a larger scale. Furthermore, such repositories include more than video (e.g., digital files of students' classroom work), and for this reason the phrase *virtual repository* is used to characterize a distributed set of heterogeneous video, metadata, client tools, and other digital resources contained in a single searchable archive. Examples of such virtual repositories are provided by YouTube and Google Video, where video files and metadata associated with them are stored and accessed across many thousands of computer servers.

A virtual repository is a key element in collaborative research because it provides a research community with an accessible touchstone corpus of empirical materials, analyses, and tools (Berman, Fox, & Hey, 2003). No such virtual repository for video data currently exists in the human sciences, although the Open Video Project has developed a large tested resource for digital video research work on such topics as automatic segmentation and summarization, and face recognition algorithms (Geisler et al., 2002). The closest analogue to a video repository currently in the United States is TalkBank, which provides a few heavily used and oft-cited data corpora (particularly audio data) in a number of language-related subdisciplines. A European video repository for educational research—Vidéo de Situations d'enseignement et d'Apprentissage (ViSA)—has recently been launched, although the videos and most information currently available are in French (ViSA, 2009). The objectives of ViSA are (a) to collect and make available to the scientific community video recordings of teaching and learning and associated documentation; and (b) to provide tools for creating, managing, and analyzing video recordings as well as for sharing of annotations and scientific productions. A repository with similar aims that is currently under development in the United States with National Science Foundation (NSF) funding is Video Mosaic (VM), which is being developed and maintained by the Rutgers University Library to house and make available the Robert B. Davis Institute for Learning video collection. This collection contains more than 3,500 hr of recordings based on developmental, longitudinal, and classroom studies of children's mathematical thinking (Maher, 2005, 2008). Eventually, VM will support uploading, sharing, and collaborative analysis of multiple video collections. This project will also address some of the metadata issues previously discussed. However, VM will require years of development and at the time this article was written, is not widely available, although a prototype with a small collection currently exists. Updates on this project can be obtained from the project Web sites (Video Mosaic Collaborative, 2009; Video Mosaic Collaborative at UW–Madison, 2009).

Although a research community might be built around a single site-specific repository, video storage requirements demand distributed storage. The storage needs are vast for even 100 researchers contributing a few hundred hours of video each (a common corpus size for a single study) at a variety of resolutions and different compression ratios. This is a "small" corpus relative to the many thousands of human scientists using video integrally in their research. A moderate-size research community would need to store and manage hundreds of terabytes (TB) of video (with petabytes and exabytes within close view). For comparative reference, consider that television worldwide generates about 31 million hr of original programming (~70,000 TB) each year, whereas worldwide production of original film is about 25,000 TB per year (Lyman & Varian, 2003). Several research centers already serve data from petabyte-size storage archives (Wikipedia, 2009).

An important research issue in distributed storage is insulating users and applications from idiosyncratic features of multiple repositories. This will require an intermediate software layer that can query each specific repository and translate whatever data are returned into a standard form. This software architecture would provide repository services to client applications via a public interface. The software would interact with repository-specific translation components that map generic calls for access, search, and retrieval into repository-specific interfaces. Using this client software, repositories could expose their contents to all members of the collaboratory without altering their practices for storing and retrieving video and metadata, so long as they also implement a version of the translation layer.

Developing a practical search function is perhaps the greatest challenge in establishing a virtual video repository. There is a large gap between the ideal of a single searchable repository and the reality of repositories with heterogeneous metadata schemes, some standardized and others ad hoc. We suggest that at least three different types of search capabilities will need to be developed, each having implications for metadata development and the functions of the software layer that translates between the generic virtual repository interface and the specific interfaces of each local video repository:

1. Full-text search of all metadata.
2. Core-metadata search. The repository would support a core set of metadata (e.g., Dublin Core[4]) guaranteed to apply to all local repositories. Thus, resources in all participating repositories would have a base level of visibility.
3. Extended-metadata search. The repository would expose to the user information about all of the metadata schemas available across the local repositories to which the repository has access. Users would select metadata schemas for searches. Only repositories supporting those schemas would be queried.

Different researchers' video repositories should not have to re-index data to a common metadata scheme; rather, all members of a collaboratory's repository must support a core metadata set (such as the Dublin Core) while exposing the user of the virtual repository to a broader range of metadata schemes. The development of a scientific field is enhanced by the common ground established by shared metadata. An example of researchers working toward this goal is Carnegie Mellon University's Brian MacWhinney (n.d.-b) and his group. They have been building into the TalkBank XML schema a system for classifying interactional structures—metadata characterizers based on the vocabulary of analytic methods such as conversational analysis, speech act theory, discourse analysis, and classical rhetoric. This metadata development is to be compliant with the Open Language Archives Community (2008) and the larger Open Archives Initiative (2008).

Attribution and Reuse

The sharing of video data among learning scientists could, with suitable human subjects protections, help accelerate the growth of scientific understanding of learning and teaching as multiple researchers gain access to video data records that now tend to reside on the shelves, on DVDs, and on hard disks of individual researchers. We recommended that researchers make original video data available to other qualified researchers and users with the provision that they agree to abide by legal and ethical guidelines governing use, reuse, and attribution. Negotiating exactly what those legal and ethical guidelines are is part of the boundary object work

[4]The Dublin Core Metadata Initiative is an open organization engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models. The Dublin Core metadata set provides a simple and standardized set of conventions for describing things online in ways that make them easier to find. Dublin Core is widely used to describe digital materials such as video, sound, image, text, and composite media such as Web pages (Dublin Core Metadata Initiative, 2009).

that remains for the community of learning sciences video researchers to accomplish. Our brief thoughts on these issues pertain to helping this community overcome barriers to broad sharing and reuse of video. We note that in some instances there are barriers to broad sharing because of concerns over the rights of human subjects who appear in the videos. Discussion of this human subjects issue is covered later.

Standards and policies for attribution and reuse. Researchers need to develop sensible attribution and authorship policies for video data. Both the metadata for and analyses of video records need to indicate authorship and attribution. There are considerable subtleties here that will require learning from best practices in related fields (e.g., motion pictures, music, photography, published written works). For example, there are now Creative Commons licensing schemes under which authors may choose to contribute their media for the public good, with or without attribution, for specific purposes (Creative Commons, 2009). The nonprofit Creative Commons licenses build on the “all rights reserved” concept of traditional copyright by offering different simple licenses that follow a voluntary “some rights reserved” approach (e.g., free reuse and sometimes remixing with attribution and nonprofit motives). The Carnegie Gallery of Teaching and Learning is a Web site that made use of this option to establish a flexible range of protections and freedoms for creators, including researchers and teachers. The rapid uptake within many communities of Creative Commons licenses that content creators use to freely assign rights to their texts, photos, music, and videos enabling their reuse and adaptation has illuminated frontiers for intellectual property rights related to educational and learning research video.

Video Research Technologies: Concluding Comment

Those who conduct video research in the learning and educational sciences are more likely to advance cumulative knowledge building if a major part of their research activity includes sharing and vetting not only their video and research findings but also the various boundary objects that are integral to the sociotechnical practices of video research. Technological tools, including new ways to harness the Internet, are central to the evolution of good practices. Returning to our introductory thought about words being chains that free us and applying these to video research, we see flexible standardized technologies as essential chains that will allow the video research community to make progress. Envisioning, designing, achieving, sharing, and widely distributing the kinds of technological boundary objects that support collaborative research with video records are important aspects of the research agenda in the learning sciences. Next we address a number of ethical issues that such collaboration raises.

ETHICAL CONCERNS IN VIDEO DATA COLLECTION

Standards for Protecting Human Subjects in Research

"Three basic principles, among those generally accepted in our cultural tradition, are particularly relevant to the ethics of research involving human subjects: the principles of respect of persons, beneficence and justice" (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979, Part B). These words from the Belmont Report, commissioned in 1974 by the National Research Act, were written in response to ethical questions raised by troubling research practices that were common during the 19th and early 20th centuries. For example, in the 1940s, the Tuskegee syphilis study used disadvantaged rural Black men to study the untreated course of a disease, depriving subjects of effective treatment in order not to interrupt the project after such treatment became available.

The standards set forth in the Belmont Report are in part implemented in the United States by "The Common Rule" (U.S. Department of Health and Human Services, 2005), regulations governing procedures for ethical treatment of human subjects in research funded by 18 U.S. federal agencies. In addition to this code, institutions, including researchers' places of employment as well as the school districts or other organizations in which they conduct research, often have requirements that regulate researchers' interactions with human subjects, ensuring that such interactions meet the requirements of the Family Educational Rights and Privacy Act (U.S. Department of Education, 2008) whether or not that research is federally funded. For example, doctoral dissertations typically must follow the regulations of the degree-granting institution, whereas classroom researchers must adhere to regulations developed by school districts. In addition, researchers are bound by codes of ethics associated with the professional societies to which they belong, such as the American Educational Research Association, the American Psychological Association, and the American Anthropological Association.

Although the complex problem of ensuring ethical, humane, dignified treatment for human participants in research cannot be fully summarized here, we briefly characterize such treatment as requiring that subjects be fully informed about the purposes, risks, and potential reward of the research; that given this information they participate voluntarily; that they be allowed to comfortably withdraw their participation during a study without penalty; and that their expectations and rights to privacy and confidentiality be honored. Furthermore, if the participant is a minor child or a member of a vulnerable population (e.g., a cognitively disabled individual), ethical treatment requires that he or she be assisted in making the decision to participate by a cognizant and responsible third party, such as a parent or legal guardian, whose major concern is the participant's welfare.

In the early stages of planning a study, the researcher should fully understand what principles and codes for the ethical treatment of human subjects apply to the proposed research. Although this may seem like a daunting assignment, there are good reasons to examine this issue carefully. First, conducting research with human subjects in accordance with principles of respect for persons, beneficence, and justice is a vital professional responsibility. Second, the principles and regulations governing ethical treatment of human subjects in research are complex, subject to interpretation and disagreement. Researchers often present their research proposals for approval to multiple institutional review boards (IRBs): at a minimum, one associated with the researcher's place of employment and one associated with the organization in which the research is conducted. When researchers collect data in multiple organizations or cross national or international lines, or if the research involves vulnerable populations, the regulatory procedures involving the protection of human subjects can become complex. Different IRBs, even boards within the same institutions that change membership over time, interpret required regulations differently in ways that impose different restrictions on different researchers at different times. Hence, disagreements between researchers and IRBs regarding the treatment of human subjects are not uncommon. Knowledge of ethical principles and regulations helps researchers prepare well-designed and well-justified protocols that also help educate IRBs regarding what is and is not acceptable within their particular sphere of research. When disagreements arise, such knowledge can be a powerful ally in making an argument with an IRB.

One common source of disputes is the time required for an IRB review. Federal regulations and the review procedures of most institutions allow for different levels of review depending on the degree of risk or inconvenience that a study imposes upon participants. However, many educational research studies should qualify for exempt status or expedited review rather than the more elaborate full IRB review that can delay the study. This is because they either study classrooms or other organizations as natural contexts without intervention or impose experimental interventions and assessments that only very minimally modify standard educational practices in which students and teachers already participate.

Another area of potential disagreement between researchers and IRBs pertains to methods for obtaining informed consent from subjects and levels of consent required. For example, if assessment or interview data can be easily codified to protect participant anonymity, and if the researcher is able to promise confidentiality based on limited access to the data that are collected, such studies present little confidentiality or privacy risk for participants and should not require elaborate informed consent procedures with signatures from both students and parents.

However, especially when minor children or vulnerable populations are subjects in research, IRBs may be cautious and impose stringent review processes and procedures. This is particularly true when video is involved. The remainder of our commentary concerns ethical issues raised by the current and future directions in video research that we have described.

Human Subjects Issues Related to Video Research

The mere addition of video collection may require in some institutions that the research proposal receive a "full review" from the IRB. Studies using video may be treated much like medical and psychological investigations for interventions for which privacy is a paramount concern. When researchers collect recordings in which individual subjects might be recognized by viewers, special concerns arise related to subject privacy and confidentiality. Participants who are recorded cannot be ensured of anonymity unless their identities are technically filtered or masked, but both filtering and masking are expensive processes that could compromise the data set for many research purposes. However, although video data are inherently non-anonymous, confidentiality can be protected in many ways, such as by restricting access to the video and to personal information such as the names of the participants or the schools in which data were collected.

Because the non-anonymous nature of video seems to make it potentially risky (e.g., it can be used in a way that might embarrass subjects), IRBs may develop special regulations and rules that apply only to video research. For example, a Web site at one university we inspected identified video educational research as being eligible for expedited review but indicated that researchers were required to adhere to the following rules of informed consent:

> It is the policy of this IRB that if the research involves the use of image or audio recording of participants, the consent form should clearly state that fact. In addition, there should be a statement about how the recordings will be used and how long they will be kept. This statement should include who will see or hear the recording and where it will be used (e.g., in a classroom, professional meeting). If the investigator wants permission for the recording to be viewed or heard by anyone other than the research staff, or if it involves sensitive material, participants should also be given an opportunity to view (or listen to) the recording after it is completed. Permission for the tape to be used should then be obtained.

This potentially highly restrictive rule—requiring researchers to obtain new consents from recorded individuals each time a recording is edited and reused outside the original research setting—would be very expensive and difficult to implement; it could limit broad sharing of video among researchers.

Issues Regarding Broad Sharing of Video Data Sources

As institutions, their IRBs, researchers, and federal agencies struggle to adequately address privacy and confidentiality issues related to video and the rights of participants in video research, it is important to acknowledge the nature of the dynamically changing social context in which this discourse is taking place. One aspect of this social context is that video recording is becoming ubiquitous in life, and social acceptance of its widespread use for many purposes is increasing. Also, there is an increased emphasis by research communities and federal agencies on the need for broad sharing of data sources within and across research communities. For example, the National Institutes of Health now requires data sharing by the researchers it funds, although video data sources may be exempt from that requirement.

The use of *data corpora* (i.e., large, amassed data collections) is common in the human sciences. A single data corpus may be used over time by multiple research groups to address many different kinds of research questions. The creation of digital libraries for research that include videos of human interaction is currently on the rise, as exemplified by the TIMSS video database (National Center for Education Statistics, 2008) used to study international instructional practices and the Child Language Data Exchange System (CHILDES; part of TalkBank) collection (MacWhinney, n.d.-a) used to study linguistic development in children. Many video research projects are supported with public funds. Hall (2000) argued that data collected on those projects should be considered a public resource. Yet what parts of a video corpus should be a public resource and for what purposes are controversial issues. A major difficulty is that video data, once collected, might be used in multiple ways, and it may not be possible at the time that a recording is made to specify precisely what these uses might be. Furthermore, often it is not known at the time of collection even who will be studying the data, because data corpora may be shared within a research community. As video records travel further from the research project in which they were collected, the types of users and uses of the video may expand in unpredictable ways, again foregrounding the importance of registering extensive metadata on the contextual aspects of video data capture and attributions for subsequent citation and source tracking. New users may have less and less knowledge about the conditions under which the video was originally collected, which creates substantial potential for inappropriate selection, use, and interpretation (Hall, 2000; Miller & Zhou, 2007). This problem is illustrated in Figure 2 (from Hall, 2000). It is therefore a matter of ethical concern to include within the corpus (not strip from it) adequate documentation about the video, including information about research subjects, so that future users of the video will be adequately informed about the nature of the video they are analyzing and how it was collected.

IRB responses to a researcher's request to maintain a fully documented video corpus indefinitely and to give wide future access to this corpus through the Internet will vary widely from one institution to another and even across time in a single institution as IRB membership changes. In some cases researchers might encounter IRB members whose scientific training orients them very directly to the

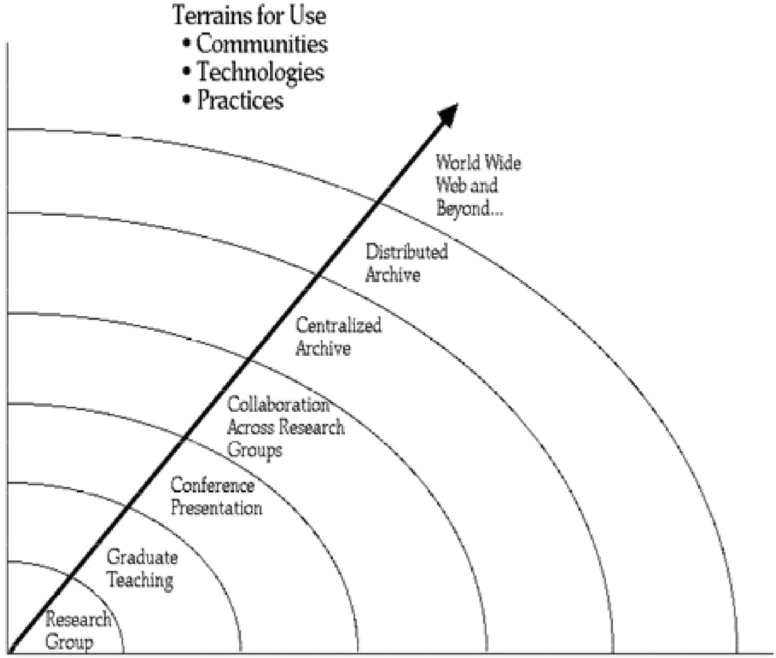

FIGURE 2 A diagram showing how contexts of use for video records are increasingly distant from the research project that produces them, making uses less predictable.

empirical testing of specific hypotheses. Hence, negotiating protocols and the conditions of informed consent that permit *emergent* and *future* analyses of video corpora will be difficult and may require substantial time and argument by the original researchers. In some cases, IRBs may request that the corpus be made anonymous or that confidential information be removed from it, or that it be allowed to exist for only a limited period of time and then be destroyed. In many instances, studies that could be conducted by graduate students or researchers who were not involved in the original data collection will be threatened because of issues of informed consent and privacy raised by IRBs. In many cases, researchers should expect that board members will require substantial education as to the nature and value of collaborative video-based educational research and the researchers' desire to collect data that can be kept in perpetuity and uploaded to shared video repositories.

Protocols and Informed Consent

Two-stage model. One way of protecting the rights of human subjects while making it possible to build and use shared repositories of video and other digital resources is to use a two-stage model involving two different types of protocols: collection protocols and use protocols.

A *collection* protocol is used to build a collection or data corpus. The collection protocol specifies how data will be collected, who will do the collecting, and what the constraints on ultimate use will be. The collection protocol must be reviewed and approved by the local IRB. The collection protocol must be renewed and kept active for the full period of data collection and initial intended use. However, the corpus itself may continue to exist after the collection protocol has been terminated. A collection protocol will almost always include preparation and use of one or more informed consent forms (ICFs).

Use protocols are always dependent on a previously approved collection protocol. Use protocols have no associated ICFs but are bound by the terms of the ICF(s) of their parent protocol. Like collection protocols, use protocols must be reviewed and approved by the IRB of the institution that owns and maintains the data corpus. An approved use protocol must be in place for the full period during which some responsible party has access to some or all data residing in the corpus for the purposes of carrying out some form of approved research. The use protocol should include the original ICF and specify what data will be accessed, who will have access, how access will be implemented and controlled, what the plans for publication are, and possibly what the plans for data destruction are when the study is completed.

Graduated model. Another approach for protecting human subjects while enabling corpus-based research with video or audio recordings is to use a graduated model. This strategy is used in TalkBank (MacWhinney, n.d.-b) and also in medical research for tissue banking. Basically, the consent form provides a menu of permissible uses (e.g., viewing by the research team only, restricted sharing among research teams, presentation at professional meetings, full Web distribution), and the subject checks off which uses he or she will permit. The graduated model could be combined with the two-stage model previously discussed, although it would be complicated to manage data collected with the combined model.

ICFs. ICFs must always fully inform participants but can describe either restrictive or permissive uses of the video. The most permissive type of ICF may ask that the subject (or the responsible parent or guardian) consent to unrestricted use of the video, including Internet use, with no financial consideration due to those recorded and no *guarantees* that the video will be used for the research purposes intended even though the researcher will make reasonable efforts to ensure appropriate use and distribution. The virtue of a permissive ICF is that it allows unrestricted sharing for appropriate research purposes and can typically be written in brief language. The disadvantages of a permissive ICF are that the wording may be questioned by an IRB. And participants, especially parents or guardians of minor children, may be especially reluctant to sign the forms. This could seriously reduce research participation.

A highly restrictive ICF, in contrast, might promise the following: very limited access, with users to include only the researcher's staff; full anonymity of videotaped subjects beyond the immediate context of the research by masking of faces and voices for any public showing or even for data sharing; permission to review and approve video releases in advance; specific dates for destruction of data; and specific ways in which the video will be protected from access by unauthorized persons, such as by the use of high-security networks and password-protected sites. The advantage of a highly restrictive ICF is that IRBs may approve the study more readily and research subjects may feel more comfortable and be more willing to participate knowing their privacy and interests are protected. The significant disadvantages are the huge increases in time and expense required for carrying out the conditions of the agreement as well as an inability to archive and share the video for research by persons beyond those directly involved in the data collection. Highly restrictive ICFs are also typically very long, which may restrict participation in some contexts. For example, Palmquist and Crowley (2007) found in their research on families who visit museums that participants were likely to decline participation if they were first required to read detailed forms.

In reality, most ICFs represent a compromise between the most restrictive and most unrestrictive cases. A rule of thumb that supports sharing is to create the least restrictive ICF possible that adequately protects subjects but also encourages broad and appropriate uses of video. In an appendix to our report to the National Science Foundation (Derry, 2007a) we included several samples of moderately permissive ICFs that have been approved by institutional IRBs in the past. Further examples and an excellent set of guidelines for developing ICFs that allow restricted use within an online video research community can be found under the section titled "The Ground Rules" at the TalkBank (MacWhinney, n.d.-b).

Ethical Issues: Concluding Comment

With increasing numbers of learning scientists and educational researchers collecting increasingly large volumes of video data that can now be shared with increasing technical ease, it becomes important to address the ethical issues associated with such sharing with the hopes of advancing scientific understandings of learning and educational practices that could be realized as benefits from such sharing. Our hope is that these comments on ethics will enable and encourage more learning scientists and educational researchers to be proactive in educating, participating in, and influencing IRBs and to develop and share protocols and ICFs that both protect subjects and enable wide sharing and future use of video data collected for educational research. We believe that many people who consent to be recorded on video for research purposes hope to benefit the wider community by doing so and that this potential benefit is increased when the video data are made available to other qualified researchers, provided those researchers agree to abide by the conditions of use approved by the original IRB.

CONDUCTING VIDEO RESEARCH: AN EPILOGUE

We conclude these guidelines by emphasizing several cross-cutting themes that repeatedly emerged throughout this project—at our conferences and meetings and across the various topics addressed as we wrote this article. These themes arise as productive tensions that help encourage new and continuing members of our video research community toward a broad middle channel that avoids the Scylla of strict and formal empiricism on the one side and the postmodern Charybdis of "anything goes" on the other. Our guidelines advocate balance at all stages in the research process between strong theory, the need for advanced planning, and formal systems of sampling and hypothesis testing on the one side, and the need to remain open and flexible to serendipitous learning, discovery, challenging of current ideas, and progressive and iterative refinement of hypotheses on the other.

There is also a balance to be sought in communities' needs and desires to collaborate around shared boundary objects that include technologies, methods, and video, thus promoting the wise reuse of precious resources and the development of strong interdisciplinary knowledge. However, such sharing raises many ethical and technical issues and practical difficulties that we must resolve together in order to sufficiently protect the rights of human subjects and our intellectual property and work as researchers. Additional tensions develop as we consider the need to come together to standardize tools and methods that will permit collaboration by learning science communities throughout the world during a time of possibility and excitement characterized by rapid new technological developments, including new cyber infrastructure initiatives (e.g., National Science Foundation Task Force on Cyberlearning, 2008).

Today's video researchers must strive to become adaptive experts with knowledge of the many issues covered in these guidelines, ranging from analysis to technology to ethics. But no one researcher can do it all. And so we conclude that the very objects of our video research, from the proposals we write to the videos and articles we publish, must include programs of work that will create infrastructures—boundary objects allowing us to become an adaptive, distributed, collaborative, expert community.

ACKNOWLEDGMENTS

The conference and report that led to this article were supported by the Data Research and Development Center, a research and technical center at the University of Chicago funded by the U.S. federal government's Interagency Education Research Initiative. The work was funded by National Science Foundation Grants

0129365 and 0015661. The first two authors co-led the conference. The remaining authors are listed in alphabetical order. We thank Kevin Brown, Barbara Schneider, and Sarah-Kathryn McDonnell for their support during this project. We thank Hilda Borko, Dan Hickey, Katherine Lewis, Kevin Miller, Rand Spiro, and Dan Schwartz for valuable input to the earlier report (Derry, 2007a), which includes a complete list of conference participants.

REFERENCES

Advanced Distributed Learning. (2004). *SCORM* (4th ed.). U.S. Office of the Under Secretary for Defense Personnel and Readiness.

Alibali, M. W., & Nathan, M. J. (2007). Teachers' gestures as a means of scaffolding students' understanding: Evidence from an early algebra lesson. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 349–365). Mahwah, NJ: Erlbaum.

Angelillo, C., Rogoff, B., & Chavajay, P. (2007). Examining shared endeavors by abstracting video coding schemes with fidelity to cases. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 189–206). Mahwah, NJ: Erlbaum.

Ash, D. (2007). Using video data to capture discontinuous science meaning making in non-school settings. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 207–226). Mahwah, NJ: Erlbaum.

Atkinson, J. M., & Heritage, J. (Eds.). (1984). *Structures of social action: Studies in conversation analysis.* Cambridge, England: Cambridge University Press.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis.* Cambridge, England: Cambridge University Press.

Barron, B. (2000). Achieving coordination in collaborative problem-solving groups. *Journal of the Learning Sciences, 8,* 403–436.

Barron, B. (2003). When smart groups fail. *Journal of the Learning Sciences, 12,* 307–359.

Barron, B. (2006). Interest and self sustained learning as catalysts for development. *Human Development, 49,* 193–224.

Bateson, M. C. (1989). *Composing a life.* New York: Atlantic Monthly Press.

Beardsley, L., Cogan-Drew, D., & Olivero, F. (2007). Video paper: Bridging research and practice for pre-service and experienced teachers. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 479–493). Mahwah, NJ: Erlbaum.

Becker, H. S. (1998). *Tricks of the trade.* Chicago: University of Chicago Press.

Berman, F., Fox, G. C., & Hey, A. J. G. (2003). (Eds.). *Grid computing: Making the global infrastructure a reality.* New York: Wiley.

Borges, J. L. (1999). On exactitude in science. In A. Hurley (Trans.), *Collected fictions* (p. 35). New York: Penguin.

Bowker, G., & Star, S. L. (2000). *Sorting things out: Classification and its consequences.* Cambridge, MA: MIT Press.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Carnegie Foundation for Teaching and Learning. (n.d.). *The gallery of teaching and learning.* Retrieved April 18, 2009, from http://gallery.carnegiefoundation.org/

Cherry, G., Fournier, J., & Stevens, R. (2003). *Using a digital video annotation tool to teach dance composition.* Retrieved April 18, 2009, from http://imej.wfu.edu/articles/2003/1/01/

Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences, 6,* 271–315.

Clancey, W. J. (in press). A transactional perspective on the practice-based science of teaching and learning. In T. Koschmann (Ed.), *Theorizing learning practice.* Mahwah, NJ: Erlbaum.

Cobb, P., & Whitenack, J. W. (1996). A method for conducting longitudinal analyses of classroom videorecordings and transcripts. *Educational Studies in Mathematics, 30,* 213–228.

Collaboration and Learning Laboratory (CaLL). (n.d.). Retrieved April 18, 2009, from http://www.siumed.edu/call/index.html

Creative Commons. (2009). Retrieved at http://creativecommons.org

Derry, S. J. (Ed.). (2007a). *Guidelines for conducting video research in education: Recommendations from an expert panel.* Chicago: Data Research and Development Center. Retrieved May 2, 2009, from http://drdc.uchicago.edu/what/video-research.html

Derry, S. J. (2007b). Video research in classroom and teacher learning: Standardize that! In R. Goldman, R. Pea, B. Barron, & S. Derry (Eds.), *Video research in the learning sciences* (pp. 305–320). Mahwah, NJ: Erlbaum.

Dublin Core Metadata Initiative. (2009). Retrieved May 2, 2009, from http://dublincore.org/

Du Bois, J. W., Schuetze-Coburn, S., Cumming, S., & Paolino, D. (1993). Outline of discourse transcription. In J. A. Edwards & M. D. Lampert (Eds.), *Talking data: Transcription and coding in discourse research* (pp. 45–89). Hillsdale, NJ: Erlbaum.

Ekman, P., Friesen, W. V., & Ellsworth, P. (1978). *Emotion in the human face: Guidelines for research and an integration of findings.* New York: Pergamon Press.

Engle, R. A., Conant, F. R., & Greeno, J. G. (2007). Progressive refinement of hypotheses in video-supported research. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 239–254). Mahwah, NJ: Erlbaum.

Erickson, F. (1977). Some approaches to inquiry in school-community ethnography. *Educational Anthropology Quarterly, 8*(2), 58–69.

Erickson, F. (1982). Audiovisual records as a primary data source. *Sociological Methods and Research, 11*(2), 213–232.

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: Macmillan.

Erickson, F. (2004). *Talk and social theory: Ecologies of speaking and listening in everyday life.* Cambridge, England: Polity Press.

Erickson, F. (2006). Definition and analysis of data from videotape: Some research procedures and their rationales. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 177–205). Mahwah, NJ: Erlbaum.

Erickson, F., & Shultz, J. (1982). *The counselor as gatekeeper: Social interaction in interviews.* New York: Academic Press.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87,* 215–251.

Geertz, C. (1973). *The interpretation of cultures.* New York: Basic Books.

Geisler, G., Marchionini, G., Wildemuth, B., Hughes, A., Yang, M., Wilkens, T., et al. (2002). *Video browsing interfaces for the open video project: Conference extended abstracts of CHI 2002* (pp. 514–515). New York: ACM Press.

GEM Exchange. (2009). *Gateway to 21st century skills.* Retrieved April 18, 2009, from http://www.thegateway.org/

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* New York: Aldine.

Goldman, R. (2007). ORION™, an online digital video data analysis tool: Changing our perspectives as an interpretive community. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 507–520). Mahwah, NJ: Erlbaum.

Goldman, R., Pea, R., Barron, B., & Derry, S. J. (Eds.). (2007). *Video research in the learning sciences.* Mahwah, NJ: Erlbaum.

Goldman-Segall, R. (n.d.). *Points of viewing children's thinking.* Retrieved April 18, 2009, from http://www.pointsofviewing.com/

Goldman-Segall, R. (1994). Challenges facing researchers using multimedia tools. *Computer Graphics, 28*(1), 48–52.

Goldman-Segall, R. (1998). *Points of viewing children's thinking: A digital ethnographer's journey.* Mahwah, NJ: Erlbaum.

Goodenough, F. L. (1928). Measuring behavior traits by means of repeated short samples. *Journal of Juvenile Research, 12,* 230–235.

Goodwin, C. (1994). Professional vision. *American Anthropologist, 96,* 606–633.

Goodwin, C. (1995). Seeing in depth. *Social Studies of Science, 25,* 237–279.

Goodwin, C. (2000). Action and embodiment within situated human interaction. *Journal of Pragmatics, 32,* 1489–1522.

Goodwin, C. (2003). Pointing as situated practice. In S. Kita (Ed.), *Pointing: Where language, culture, and cognition meet* (pp. 217–242). Mahwah, NJ: Erlbaum.

Goodwin, C. (n.d.). *Charles Goodwin.* Retrieved April 18, 2009, from http://www.sscnet.ucla.edu/clic/cgoodwin/projects.htm

Gore, A. (2007). *The assault on reason.* New York: Penguin Press.

Gottman, J. M., & Notarius, C. (2002). Marital research in the 20th century and a research agenda for the 21st century. *Family Process, 41,* 159–197.

Hall, R. (2000). Video recording as theory. In A. Kelley & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 647–664). Mahwah, NJ: Erlbaum.

Hammer, D., & van Zee, E. (2006). *Seeing the science in children's thinking: Case studies of student inquiry in physical science, a staff developer's guide.* Portsmouth, NH: Heinemann.

Harrison, B. L., & Baecker, R. M. (1992). Designing video annotation and analysis systems. In K. S. Booth & A. Fournier (Eds.), *Proceedings of the Conference on Graphics Interface '92, Vancouver, BC, Canada* (pp. 157–166). San Francisco: Morgan Kaufmann.

Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences, 4*(1), 39–103.

Journal of Visualized Experiments. (2008). Retrieved April 18, 2009, from http://www.jove.com/

Koschmann, T. (Ed.). (1999). Meaning making [Special issue]. *Discourse Processes, 27*(2).

Koschmann, T., Glenn, P., & Conlee, M. (1999). Theory presentation and assessment in a problem-based learning group. *Discourse Processes, 27,* 119–133.

Lapadat, J. C., & Lindsay, A. C. (1999). Transcription in research and practice: From standardization of technique to interpretive positionings. *Qualitative Inquiry, 5*(1), 64–86.

Leacock, R. (1973). Technology and reality at the movies. *Technological Review, 75*(4), 2–7.

Lehrer, R., & Schauble, L. (2004). Modeling natural variation through distribution. *American Educational Research Journal, 41,* 635–679.

Lemke, J. L. (2000). Across the scales of time: Artifacts, activities, and meanings in ecosocial systems. *Mind, Culture and Activity, 7*(4), 273–290.

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research on Science Teaching, 38*(3), 296–316.

Leonard, M. J. (2006). *"What's the science behind it?" Models and modeling in a design for science curriculum.* Unpublished doctoral dissertation, University of Wisconsin–Madison.

Leonard, M. J., & Derry, S. J. (2006). Tensions and tradeoffs in a "design for science" classroom: The "forces in balloon lecture." In S. A. Barab, K. E. Hay, & D. T. Hickey (Eds.), *Proceedings of the Seventh International Conference of the Learning Sciences* (pp. 411–417). Mahwah, NJ: Erlbaum.

Lewis, C. (2002). *Lesson study: A handbook of teacher-led instructional change.* Philadelphia: Research for Better Schools.

Lewis, M., & Rosenblum, L. (1974). *The effect of the infant on its caregiver.* New York: Wiley.

Lu, J., & Rose, R. (2003). Seeing math through multimedia case studies. *Concord, 7*(1), 1, 4–5.

Lyman, P., & Varian, H. R. (2003). *How much information 2003?* Retrieved May 2, 2009, from http://www.sims.berkeley.edu/how-much-info-2003

Lynch, S., Kuipers, J., & Pyke, C. (2005). Examining the effects of a highly rated science curriculum on diverse students: Results from a planning grant. *Journal of Research in Science Teaching, 42,* 912–946.

Mackay, W., & Beaudouin-Lafon, M. (1998). DIVA: Exploratory data analysis with multimedia streams. In *CHI 1998 Proceedings* (pp. 416–423). New York: ACM Press.

MacWhinney, B. (n.d.-a). *Child language data exchange system.* Retrieved April 18, 2009, from http://childes.psy.cmu.edu/

MacWhinney, (n.d.-b). *TalkBank.* Retrieved April 18, 2009, from http://talkbank.org/

Maher, C. (2005). How students structure their investigations and learn mathematics: Insights from a long-term study. *Journal of Mathematical Behavior, 24,* 1–14.

Maher, C. (2008). Video recordings as pedagogical tools in mathematics teacher education. In D. Tirosh & T. Wood (Eds.), *Tools and processes in mathematics teacher education* (pp. 65–83). Rotterdam, The Netherlands: SensePublishers.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom.* Cambridge, MA: Harvard University Press.

Miller, K. F., & Zhou, X. (2007). Learning from classroom video: What makes it compelling and what makes it hard. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 321–334). Mahwah, NJ: Erlbaum.

Mischler, E. G. (1991). Representing discourse: The rhetoric of transcription. *Journal of Narrative and Life History, 1,* 255–280.

National Center for Education Statistics. (2008). *Trends in international mathematics and science study.* Retrieved April 18, 2009, from http://nces.ed.gov/timss/

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for protection of human subjects of research.* Retrieved May 2, 2009, from http://ohsr.od.nih.gov/guidelines/belmont.html

National Science Foundation Task Force on Cyberlearning. (2008). *Fostering learning in the networked world: The cyberlearning opportunity and challenge.* Retrieved May 2, 2009, from http://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf08204

Nemirovsky, R., Lara-Meloy, T., Earnest, D., & Ribeiro, B. T. (2001, July). *Videopapers: Investigating new multimedia genres to foster the interweaving of research and teaching.* Paper presented at the 25th Meeting of the International Group for the Psychology of Mathematics Education, Utrecht University, The Netherlands.

Ochs, E. (1979). Transcription as theory. In B. Schieffelin (Ed.), *Developmental pragmatics* (pp. 43–72). New York: Academic Press.

Ochs, E., & Taylor, C. (1996). The 'father knows best' dynamic in family dinner narratives. In K. Hall (Ed.), *Gender articulated: Language and the socially constructed self* (pp. 97–121). New York: Routledge.

Open Archives Initiative. (2008). Retrieved April 18, 2009, from http://www.openarchives.org/

Open Language Archives Community. (2008). Retrieved April 18, 2009, from http://www.language-archives.org/

Palmquist, S. D., & Crowley, K. (2007). Studying dinosaur learning on an island of expertise. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 271–286). Mahwah, NJ: Erlbaum.

Parten, M. B. (1932). Social participation among preschool children. *Journal of Abnormal and Social Psychology, 27,* 243–269.

Pea, R. D. (1999). New media communication forums for improving education research and practice. In E. C. Lagemann & L. S. Shulman (Eds.), *Issues in education research: Problems and possibilities* (pp. 336–370). San Francisco: Jossey-Bass.

Pea, R. D. (2006). Video-as-data and digital video manipulation techniques for transforming learning sciences research, education and other cultural practices. In J. Weiss, J. Nolan, & P. Trifonas (Eds.), *International handbook of virtual learning environments* (pp. 1321–1393). Dordrecht, The Netherlands: Kluwer Academic.

Pea, R. D., & Hay, K. (2003, March). *Report to the National Science Foundation: CILT Workshop on Digital Video Inquiry in Learning and Education, November 25–26, 2002* (based on NSF #0124012). Stanford, CA: Stanford Center for Innovations in Learning.

Pea, R., & Hoffert, E. (2007). Video workflow in the learning sciences: Prospects of emerging technologies for augmenting work practices. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 427–460). Mahwah, NJ: Erlbaum.

Pea, R., & Lindgren, R. (2008). Video collaboratories for research and education: An analysis of collaboration design patterns. *IEEE Transactions on Learning Technologies, 1*(4), 235–247.

Powell, A. B., Francisco, J. M., & Maher, C. A. (2003). An analytical model for studying the development of learners' mathematical ideas and reasoning using videotape data. *Journal of Mathematical Behavior, 22,* 405–435.

Roschelle, J., Pea, R. D., & Trigg, R. (1990). *VideoNoter: A tool for exploratory video analysis* (Tech. Rep. No. 17). Palo Alto, CA: Institute for Research on Learning.

Schatzman, L., & Strauss, A. L. (1973). *Field research: Strategies for a natural sociology* (pp. 94–107). New Jersey: Prentice Hall.

Schegloff, E. A. (1997). Whose text? Whose context? *Discourse & Society, 8*(2), 165–187.

SciVee. (2009). Retrieved April 18, 2009, from http://www.scivee.tv/

Sfard, A., & McClain, K. (2002). Analyzing tools: Perspectives on the role of designed artifacts in mathematics learning [Special issue]. *Journal of the Learning Sciences, 11*(2/3).

Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1–27). New York: Elsevier Science.

Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 383–395). Mahwah, NJ: Erlbaum.

Sherin, M. G., & Sherin, B. L. (2007). Research on how people learn with and from video. In S. Derry (Ed.), *Guidelines for conducting video research in education: Recommendations from an expert panel* (pp. 47–58). Chicago: Data Research and Development Center. Retrieved May 2, 2009, from http://drdc.uchicago.edu/what/video-research-guidelines.pdf

Star, S. L. (1989). The structure of ill-structured solutions: Boundary objects and heterogeneous distributed problem solving. In L. Gasser & M. Huhns (Eds.), *Distributed artificial intelligence, Vol. II* (pp. 37–54). London: Pitman.

Stern, D. (1977). *The first relationship.* Cambridge, MA: Harvard University Press.

Stevens, R. (2007). Capturing ideas in digital things: A new twist on the old problem of inert knowledge. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 547–563). Mahwah, NJ: Erlbaum.

Stevens, R., & Hall, R. (1998). Disciplined perception: Learning to see in technoscience. In M. Lampert & M. L. Blunk (Eds.), *Talking mathematics in school: Studies of teaching and learning* (pp. 107–150). New York: Cambridge University Press.

Stevens, R., Satwicz, T., & McCarthy, L. (2008). In-game, in-room, in-world: Reconnecting video game play to the rest of kids' lives. In K. Salen (Ed.), *The ecology of games: Connecting youth, games and learning* (pp. 41–66). Cambridge, MA: MIT Press.

Stevens, R., & Toro-Martell, S. (2003). Leaving a trace: Supporting museum visitor interpretation and interaction with digital media annotation systems. *Journal of Museum Education, 28*(2), 25–31.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap: Best ideas from the world's teachers for improving education in the classroom.* New York: Free Press.

Teachscape. (2009). Retrieved April 18, 2009, from http://www.teachscape.com/

Tobin, J. J., & Hsueh, Y. (2007). The poetics and pleasures of video ethnography of education. In R. Goldman, R. Pea, B. Barron, & S. Derry (Eds.), *Video research in the learning sciences* (pp. 77–92). Mahwah, NJ: Erlbaum.

Tobin, J. J., Wu, D. Y. H., & Davidson, D. (1989). *Preschool in three cultures: Japan, China, and the United States.* New Haven, CT: Yale University Press.

Transana. (2005). *Qualitative analysis software for video and audio data.* Retrieved April 17, 2009, from http://www.transana.org

Trevarthen, C., & Aitken, K. J. (2001). Infant intersubjectivity: Research, theory, and clinical applications. *Journal of Child Psychology and Psychiatry, 42,* 3–48.

U.S. Department of Education. (2008). *34 CFR Part 99 family educational rights and privacy; final rule.* Retrieved April 17, 2009, from http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html

U.S. Department of Health and Human Services. (2005). *Code of Federal Regulations, Title 45 Part 46, protection of human subjects.* Retrieved May 2, 2009, from http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm

Video Mosaic Collaborative. (2009). Retrieved December 15, 2009 from http://www.video-mosaic.org

Video Mosaic Collaborative at UW–Madison. (2009). Retrieved December 15, 2009 from VMC.wceruw.org

ViSA. (2009). *Vidéo de Situations d'enseignement et d'Apprentissage (ViSA).* Retrieved May 2, 2009, from http://visa.inrp.fr

Wikipedia. (2009). *Petabyte.* Retrieved April 18, 2009, from http://en.wikipedia.org/wiki/Petabyte

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan.

Wortham, S. (2004). From good student to outcast: The emergence of a classroom identity. *Ethos, 32*(2), 164–187.

Zacks, R., & Tversky, D. (2001). Event structure in perception and conception. *Psychological Bulletin, 127,* 3–21.

Znaniecki, F. (1934). *The method of sociology.* New York: Rinehard.

Appendix A
Recording: Practical Techniques From Established Field Methods

Heuristics for Collecting Field Recordings

1. *Conduct fieldwork strategically but adapt plans* as fieldwork unfolds.
2. *Capture as much as possible, and index recordings* so events in a growing collection can be found later.
3. *Separate observations from inferences* in field notes by distinguishing between *what* is observed and *how* it unfolded (observational notes) on one hand versus *why* it happened or *what it means* (analytic notes) on the other.

Selecting Equipment

1. *Cameras.* Consider newer high-definition cameras that capture directly to memory cards/sticks/drives (e.g., Canon HG20 Vixia HD, Flip MinoHD, Kodak Zi6, Panasonic HDC-HS100). The basic mini-DV camera should have
   a. Image stabilization (for hand-held use);
   b. A microphone jack and true stereo (left/right) audio separation;
   c. A rechargeable, long-life battery;
   d. A good-quality detachable wide-angle lens;
   e. If using tapes (still a preferred medium for many), a camera into which tapes can be loaded while mounted on a tripod.

2. *Microphones.* Purchase high-quality external battery–powered microphones that send signals to battery-powered receivers that plug into the camera. Several types are useful:
   a. *Directional* or *shotgun* microphones can be mounted on a camera and selectively capture sound from where the microphone points.
   b. *Boundary* or *pressure-zone* microphones placed in a fixed location unselectively capture a wide spectrum of sound.
   c. *Lavaliere microphones* are usually pinned to a speaker or moving object close to sound sources the researcher wants to capture.
3. *Tripod, earphones, gaffer's tape, and camera bag.*
   a. A tripod enables the researcher to smoothly pan and zoom (but see [c] below) and lock the camera in a fixed position when busy with other things.
   b. Earphones should be used to check sound at the beginning and periodically while recording.
   c. Gaffer's tape (black cloth tape) will allow taping down the legs of the tripod without leaving marks on furniture or other objects.
   d. A durable, spacious camera bag will allow one person to carry the entire rig in one hand while carrying a tripod in the other. This kind of rig can be treated as a unit, and as more units are taken into the field, options for capturing multiple video and audio streams increase.
4. *Beyond the basics.* Sophisticated, expensive specialized equipment is available. For example, Koschmann uses head-mounted and fiber-optic cameras to capture surgical procedures (Collaboration and Learning Laboratory [CaLL], n.d.).

Camera Work

1. *Single camera.* The most common setup in studies of classrooms or workplaces is a single consumer-grade video camcorder. A single camera recording cannot capture much complexity, so recording necessarily shifts among activities with different structures (e.g., group presentations, whole-class discussion). The researcher should record continuously and use panning and zooming minimally. Too much selection at recording time may rule out later lines of analysis.
2. *Multiple cameras.* With two camera rigs, the researcher can distinguish between "wide" and "follow" perspectives and use left/right separable stereo sound to capture four separable audio sources. This provides relatively large capability and flexibility for capturing complex field situations. Allocating two video and four audio sources is a sampling decision (Hall, 2000). If there is a strong theoretical rationale, these sources can be allocated precisely. If a researcher is less sure of the phenomena of interest, it is best to allocate resources to maximize the chances of finding interesting things as they emerge during recording or in

later analyses. Sampling strategies for a two-camera setup are detailed in the original National Science Foundation report (Derry, 2007a).

3. *Panning and zooming.* The researcher should not focus only on part of an integrated social activity. The follow camera should stay with the *proxemic shape* of the interacting group (i.e., bodies in relation to one another and things). For example, a follow camera would attempt to have speaker *and* listeners in view as a speaker is making a point. But if the speaker begins to point to a white board or sheet of paper, the follow camera can alternate between zooming in to get the artifact details and zooming back out to capture the group. As illustrated in Figure A-1, zooming in and out of the scene is preferable to panning across speakers and media. Row 1 illustrates these two different recording strategies. To the left in Row 1, the operator is panning with tight zoom to follow sequential talk. This produces the video record in Row 2, which loses interactional context. To the right in Row 1, a skilled operator stays at wide zoom when conversation is under way, zooming in tight to capture gestures by Speaker 1 that are coupled to documents but returning to wide zoom as the conversation continues with Speaker 2. The video record shown in Row 3 preserves interactional context but still allows an analyst to recover what is contained in the documents.

Strategies for Indexing Recordings for Later Analysis

1. *Invent index structures* so analysts can find particular events without having to review the entire collection of field materials (see Schatzman & Strauss, 1973).
2. *Make time-indexed notes* as recording is in progress (e.g., 09:43 followed by a description of an event noticed). These notes can be collected in a searchable database for efficient retrieval of corresponding video segments *provided* the researcher
   a. starts multiple recordings simultaneously;
   b. includes a reference event when starting multiple devices (e.g., a clap);
   c. makes continuous recordings (the clock time written on field notes will function as an index to video only if the video record is continuous);
   d. time-indexes field notes in coordination with recordings in progress.
3. *Think about likely contexts of analysis* that will occur later. For example, the researcher might imagine sitting at a computer projecting video onto a screen, working on field materials with collaborators. The researcher would want to be able to find and play a video episode of some event of interest for discussion and analysis in a project meeting.

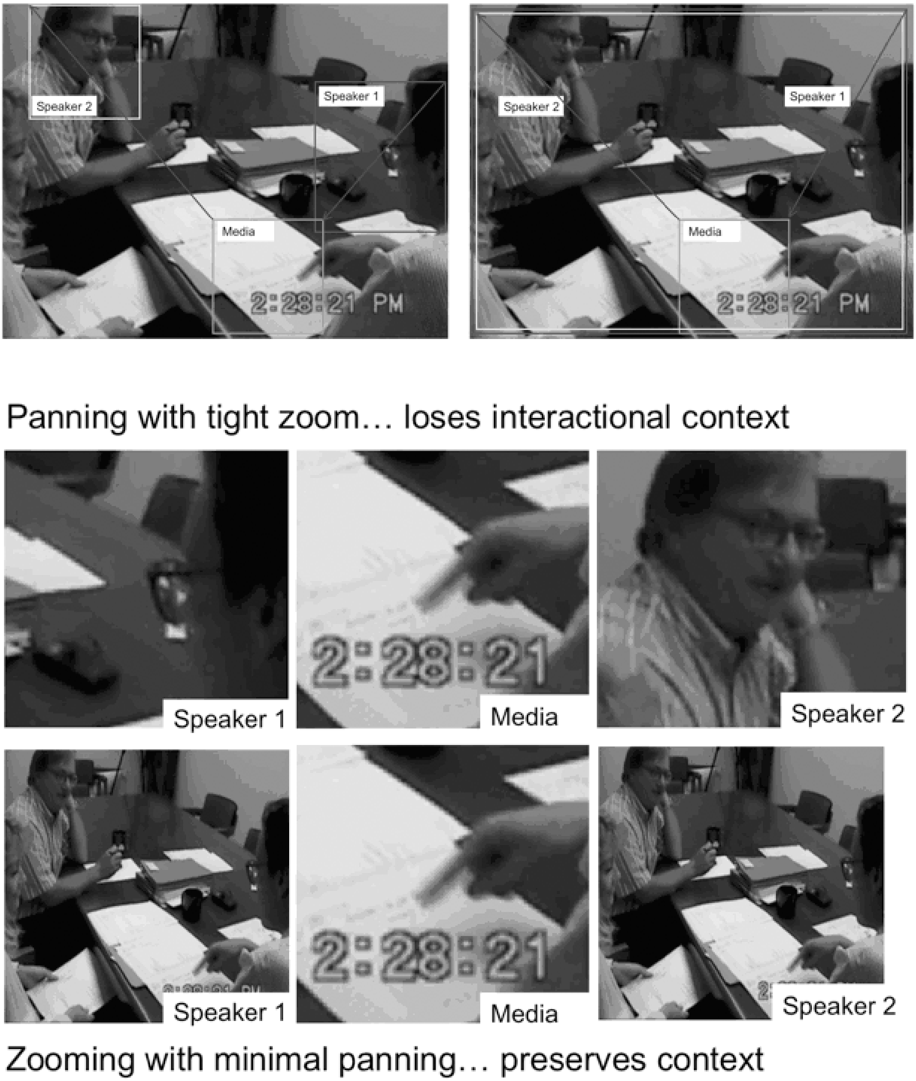

FIGURE A-1 Row 1 shows two frames from a stationary wide-angle camera. Rows 2 and 3 contrast two different recording strategies for the same time period by a follow camera.

Time Schedule

1. *Plan a feasible analysis-design cycle.* Contrary to some accounts in the literature, video recordings probably cannot be understood well enough to be immediately useful for design purposes the same or very next day.

Appendix B
Technological Tools for Video Research

Video Analysis Systems: Some Examples

1. *Orion,* formerly Constellations (Goldman-Segall, 1994), enables researchers to add and share videos and related descriptors, links, and transcripts. Users can extract clips (or "stars") from video records in a searchable database and can organize these stars into "constellations," creating sharable video narratives. Orion also supports collaborative distributed coding of video selections. Orion can be accessed free online, and an upgrade is being tested with selected users (see Goldman, 2007).
2. *Transana* (2005) allows researchers to create large video collections, organize video clips into categories, apply searchable analytic keywords to clips, engage in data mining and limited hypothesis testing across large collections, and share analytic markup with distant colleagues. It provides excellent support for transcription and supplies a number of reporting formats. Transana is a stable and inexpensive open source system with a large user community that has been adapted for varied uses, including serving as a platform for video games and simulations.
3. *WebDIVER* is a Web services environment supporting collaborative analysis that many teams use for distributed access, annotation, and search of video uploaded via a Web browser and streamed from remote servers (Pea, 2006; Pea & Lindgren, 2008). Users create perspectives by making "dives" into one or more streaming videos. As a video streams, a dive is created when the user records one or more video segments (or zooms and pans into specific regions in videos) using a virtual camera controlled by a mouse. Each dive segment is represented as a worksheet panel with an image thumbnail. Narratives, transcripts, and codes can be created and exported.

4. *VideoTraces* enables users to lay down a reflective "trace" on top of a video record. The trace consists of voice and text annotation and a pointing gesture implemented as a hand cursor. A VideoTraces file can be replayed at variable speeds to experience the overlay. VideoTraces has supported research and instruction in science education museums and higher education courses, such as rowing and dance composition (Cherry, Fournier, & Stevens, 2003; Stevens, 2007; Stevens & Toro-Martell, 2003).
5. *Proprietary software* can be distinguished from previously presented tools that were developed by researchers to support their own research but that may be shared at low cost, often informally. Proprietary video analysis products have been developed for sports and other markets (e.g., StudioCode). Proprietary software for qualitative social science research, such as NUD*IST, NVivo, and ATLAS.ti, possess capabilities for supporting video analyses, especially if based on written transcripts of video. Video editing and chunking are supported by commercial tools such as Apple's iMovie or Adobe Premiere, but these are not oriented to coding or reflection.
6. *Tools for developing and sharing video cases* are oriented toward creating and sharing video cases for professional development, although some analytical functions may be included. Examples include Teachscape (2009), in which developers make available video cases of teaching practice, and the Carnegie Foundation's Knowledge Exchange Exhibition and Presentation (KEEP) toolkit designed to help teachers share cases of their scholarship on the World Wide Web. (Although historically important, the KEEP toolkit is no longer in service.)

Example Metadata Schemes

Examples of the type of work on coding metadata that video researchers in the learning sciences must undertake include the following:

1. *The Gateway to Educational Materials* (GEM Exchange, 2009) instructional topics hierarchy builds on the Dublin Core (Dublin Core Metadata Initiative, 2009).
2. *The Open Language Archives Community* (2009) conforms to the larger Open Archives Initiative (2008). The stated goal of these initiatives is that any user on the Internet should be able to go to a single gateway to find all the relevant language resources available at all participating institutions, whether the resources are data, tools, or advice. The community ensures ongoing interoperation and quality through standards and review processes.

3. *The Sharable Content Object Reference Model* (SCORM; Advanced Distributed Learning, 2004), initiated in 1997 by the U.S. Department of Defense, is now a widely used XML-based standard and specification for Web-based e-learning objects.