

Learning from Digital Video: An Exploration of How Interactions Affect Outcomes

Robb Lindgren, Roy Pea, Sarah Lewis, and Joe Rosen
Stanford University, 450 Serra Mall, Stanford, CA 94305
Email: {robblind, roypea, sarahl, joro}@stanford.edu

Abstract: The sinking costs of producing digital video and its growing presence on the Internet suggest that it has potential for use in web-based learning technologies. However, there have been few investigations into how the kind of interaction one has with video impacts subsequent learning. In this in-progress study participants are asked to watch video of an expert taking apart a toaster and describing how it works. The recorded event is the same for all participants, but the event is presented in one of three different modalities: (1) digital video shot from a free-standing camera (2) digital video shot from a free-standing camera that has been annotated in a video mark-up application called DIVER and (3) digital video shot from a head-mounted camera. A number of different assessment tasks are used to characterize the quantity and type of learning that is supported by a particular mode of video interaction.

Video-Based Learning and Interactivity

Decreasing costs and increasing accessibility of digital video technologies has made it remarkably easy to capture and share a significant event. There is tremendous potential for people to learn from watching a video that someone else has recorded because they may contain vivid descriptions of expert knowledge or a demonstration of expert practice (e.g., a video of an art historian deconstructing the scenes in Picasso's *Guernica*). In certain visually-based domains a video is likely to have more instructional value than a text-based description of the same event. Video-sharing platforms such as *YouTube* have led to an explosion in the number of digital videos available for viewing online by a widespread and continually increasing audience. While not all of the videos uploaded on the Internet have educational value, there is clearly the potential for online video to serve as a powerful medium for learning and collaboration. For this potential to be realized, however, there is still much more that needs to be understood about how the construction and presentation of video affects the way we learn one another.

Previous research on learning from video has been limited to linear, non-interactive video, such as television (for a review see Seels, Fullerton, Berry & Horn, 1996). Digital video creates the possibility for new kinds of interactions with video, both in terms how the video is recorded and how people are able to manipulate and share existing footage. Lightweight and extremely portable video cameras can be purchased cheaply such that multiple cameras can be used to record the same event from different perspectives. Likewise, there are online tools becoming available that allow users create text annotations of video clips as well as perform basic video editing tasks. One such tool that was developed in our lab is called DIVER (Digital Interactive Video Exploration and Reflection) (Pea et al., 2004) and it was designed specifically to support online collaboration for the analysis of video data—it's a process we refer to as computer-supported collaborative video analysis, or CSCVA (Pea, Lindgren, & Rosen, 2006). While the different kinds of interactions one is able to have with digital video is rapidly increasing, there is a lack of research that speaks to how these interactions facilitate learning or create successful collaborations. In this paper we argue that this research is necessary because different ways of looking at the same event can change how one learns from that event, and thus it may be important to the design of video-based learning technologies. To this end, we describe the design of study in progress that looks at how the camera point of view affects the outcomes of a learning assessment in a novel domain.

Point of View

One way to change a learning interaction is to switch the perspective from which a learner is experiencing the event of interest. Most educational videos, for example, are recorded from the third-person perspective—a camera in a fixed position records the scene from the perspective of an “unseen other.” By contrast, current technology makes it possible to capture a scene from the first-person perspective—the camera takes on the perspective of someone actually experiencing the activity of interest. This is generally accomplished via a wearable camera that is attached to a person in such a way that the camera's field of view approximates that of the person wearing it.

There is reason to believe that viewing a digital video taken from the first person perspective would have a learning benefit over a video taken from the third-person perspective. Goodwin (1994) describes a characteristic of experts in a given domain called *professional vision*, which is a way to describe how an expert literally sees the domain differently than a novice looking at the same domain. For example, Goodwin describes how a seasoned archeologist can perceive patterns in the dirt at the site of a dig that an archeological student does not perceive. Following this logic, it's possible that if a novice had the opportunity to see the scene in the same way that an expert sees it (e.g. attending to the aspects of the scene that the expert has come to recognize as important) then it may be easier for the novice to learn in the domain and hasten the path to becoming an expert themselves. Thus, a video captured in the first-person may have a learning benefit because it contains elements of perceptual expertise that are not present in a video captured from a fixed third-person perspective.

Another reason that the first person perspective may benefit learning is that seeing a scene through the eyes of a knowledgeable actor may lead the viewer to embody the expertise demonstrated in the video. In other words, seeing an activity as an expert sees it may elicit feelings of confidence and motivation as if they were the expert themselves. A first-person perspective may also give viewers a feeling of social presence, or the sense that they are actually interacting in the environment as opposed to being a passive observer. This too may have motivational effects that lead to greater subsequent learning.

Study Design

This study was designed to contrast learning from digital video of the same event recorded in the first or third person perspective. The first-person perspective was accomplished using a camera attached to a headband worn by the subject matter expert in the video. The third-person perspective was accomplished by recording the same simultaneous event using a camera on a tripod focused over the shoulder of the subject matter expert. The specific content that we chose for this study is a video recording of an expert disassembling a toaster and describing how it works. Toasters are surprisingly complex instruments that employ several key concepts from physics and electricity. Very few people understand the functionality of a toaster, and describing how one works is a highly visual task that was well suited for a video-learning study.

Head Cam

In the last few years some researchers in the social sciences have begun equipping their participants with head-mounted cameras as a way to collect data about aspects of human behavior such as communication and collaboration (for an example, see Fussell, Setlock, & Kraut, 2003). We feel that viewing video recorded by a head cam is a unique opportunity to take on the perspective of another individual, and we were interested in how easily the adoption of an expert's perspective would translate to learning. For this study we fitted the toaster expert with a Sony Super HAD high-resolution camera that recorded to a mini-DV deck. This video was transferred to a computer and processed into a QuickTime movie file.

Study Procedures

When completed, a total of 30 participants will be run in this study with 15 participants each in the first and third-person perspective conditions. Each session begins with the participant completing a short survey that is used to assess the participant's prior knowledge of toasters and other related domains (e.g., physical mechanics). Next the participant is told that they will have the opportunity to view some video and that their task is to use the materials to learn as much as possible about how a toaster works in 15 minutes. Participants will view either the video recorded from the tripod-mounted camera or the head cam video (see Figures 1a and 1b for a screen shot of each condition):

1. *Video from a tripod-mounted camera + passage (Figure 1a)*: In this condition, participants are presented with the toaster video in the QuickTime digital video player on a laptop computer. The participant is also given a piece of paper with a passage that describes the functionality of a toaster. Some of the information in the passage is redundant information presented in the video, and some of the information is novel. Prior to starting, the participant is given brief instructions on how to use the video player.
2. *Video from a head-mounted camera + passage (Figure 1b)*: This condition is identical to the first condition except that the video used is recorded from a head camera. The head cam video was recorded at the same time as the tripod-mounted camera so that the audio and the events captured in both conditions are the same.

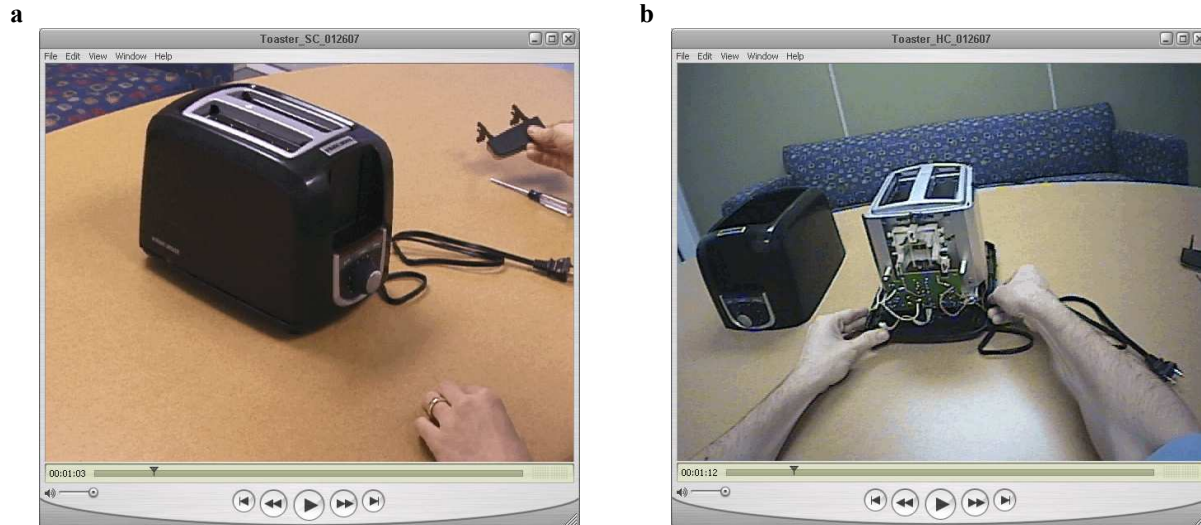


Figure 1. Toaster video from a tripod-mounted camera (a) and from a head-mounted camera (b).

Learning Assessments and Data Collection

After the participants in each of the conditions have reviewed their materials, a series of assessments are administered. Each of these assessments target a different form of knowledge or understanding in order to determine if different types of video interactions support different kinds of learning. Participants are asked to complete four tasks: (1) Answer a series of paper-based conceptual and vocabulary questions (e.g., “Why are the filaments in a toaster wound more tightly at the bottom than at the top?”) (2) Describe how a toaster works to someone who doesn’t know (3) Describe how they would troubleshoot a series of hypothetical toaster malfunctions (e.g., “The toaster is plugged in and the tray stays down, but my bread isn’t heating up.”) and (4) Draw a functional diagram of a toaster.

In addition to the paper-based learning assessments, we are also looking at the participants’ physiological arousal while watching the two different videos. Prior to watching the video the experimenter hooked the participants up to a device that records the participants skin conductance and heart rate, which are commonly used as measures of arousal.

References

- Fussell, S. R., Setlock, L. D. & Kraut, R. E. (2003). Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In: *Proceedings of the CHI 2003 Conference on Human Factors in Computing Systems*. ACM Press.
- Goodwin, C. (1994). Professional vision. *American Anthropologist*, 96, 606-633.
- Pea, R., Lindgren, R., & Rosen, J. (2006). Computer-supported collaborative video analysis. *Proceedings of the Seventh International Conference of the Learning Sciences (ICLS)*. Bloomington, Indiana. Mahwah, NJ: Erlbaum.
- Pea, R., Mills, M., Rosen, J., Dauber, K., Effelsberg, W., & Hoffert, E. (2004). The DIVER™ project: Interactive digital video repurposing. *IEEE Multimedia*, 11(1), 54-61.
- Seels, B., Fullerton, K., Berry, L., & Horn, L. J. (2004). Research on learning from television. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology*, 2nd edition. Mahwah, NJ: Erlbaum.