Investigating Theoretical Frameworks and Strategies for Instructional Video

Lara Bove
George Mason University
Fairfax, VA

Video is increasingly becoming a mainstay in education and workplace training. In 2020, 74% of corporate training was delivered through video (Bouchrika, 2020). Despite this widespread usage, it is not clear when video instruction is the best media to achieve the desired behavioral and performance outcomes. For instance, videos do not consistently support learners in being able to recall information (Chen & Thomas, 2020; van der Meij, 2019). Researchers have been investigating all different aspects of the use of video for instruction to better understand what works, what doesn’t work, and why. This understanding can help instructional designers when they are making decisions about including video as part of an overall instructional strategy.

The task for instructional designers is to select the video approach and technical features that are best suited to the instructional goals. However, the instructional designer has to consider many different aspects of the video that extend beyond the framing of the content (e.g., lecture, story, demonstration). There are developmental decisions about video length, perspective (i.e., first-person or third-person point of view), level of realism, and so much more. Throughout this paper, I will use the term “design decisions” to refer to the depth and breadth of Instructional Design decisions, which goes beyond the D in ADDIE. I begin with a literature review and an exploration of the current understanding of instructional video in practice.

Literature Review

In order to investigate the different uses of instructional video, it is important to define instructional video. In the broadest sense an instructional video is any moving picture that is used to provide instruction. Using a video as part of instruction does not make it an instructional video. If we think of pictures in a textbook, some are instructional and others are not. This could become complicated, but for purposes of this research, the video must provide instruction on its own or be critical to the instruction.

I conducted a literature review of peer-reviewed journal articles describing research into aspects of instructional video. I specifically wanted to understand what research has uncovered about different approaches using video for instructional purposes. I included studies in K-12, the university, medical schools, and the workplace, as long as the studies looked at aspects of video for instruction. I excluded studies that looked at uses and features for virtual meetings, virtual instructor led training, alternate reality/virtual reality, and even the use of video for data collection about training.

I began my literature review by looking at a 10-year period, but quickly realized that the technology was changing so quickly that a study from 10 years ago might not be relevant today. For example, van der Meij and van der Meij (2016) wanted to investigate the impact of review videos. In order to isolate the effect of the review video from the effect of rewatching a video, they removed the ability for learners to rewind or fast-forward. While this made sense with their research purposes, it demonstrated the impact of the pace of technological change. Today, most
learners watch videos from an individual device and have controls over the player. The study was only 5 years old, yet today’s learners have access to all the video controls, particularly as they watch videos on their own devices. For this reason, I restricted the literature to 5 years.

My literature review included articles from: *Journal of Workplace Learning, International Journal of Training and Development, British Journal of Education Technology, Multimedia Tools and Applications, Human Resource Development Quarterly, Educational Technology Research and Development, Performance Improvement Quarterly*. I had a total of 29 articles, and all but one involved an experiment or study of some kind. Most of the studies took place in an educational setting. Table 1 provides a count of the studies based upon different demographic information, including the location of the study.

Table 1
Demographics from Experimental and Quasi-Experimental Studies Included in Review

<table>
<thead>
<tr>
<th>Ages of Participants (years)</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>8-18</td>
<td>5</td>
</tr>
<tr>
<td>18-30</td>
<td>13</td>
</tr>
<tr>
<td>17 and older</td>
<td>4</td>
</tr>
<tr>
<td>Adults (ages not specified)</td>
<td>6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Location of study</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>School (K-12)</td>
<td>6</td>
</tr>
<tr>
<td>College or University</td>
<td>17 *</td>
</tr>
<tr>
<td>MOOC</td>
<td>1</td>
</tr>
<tr>
<td>Medical School</td>
<td>1</td>
</tr>
<tr>
<td>Workplace</td>
<td>2 *</td>
</tr>
<tr>
<td>General public</td>
<td>2</td>
</tr>
</tbody>
</table>

*Note. *Denotes inclusion of a study that is listed in both categories.

The age groupings provide a way to identify whether a study looked at school-age children, college-age adults, and larger ranges of adult ages. If a study did not provide ages, but specified adults, including college students, the study is listed under adults (ages not specified). The category of 17 and older is specifically for studies that indicated the age of adult participants. For example, Ramlatchan and Watson (2020) had participants aged 17-66 years old.

Location refers to where the study took place or, in the case of MOOC, where the video instruction was delivered. College or University includes vocational education, but MOOC is listed separately. Medical school is listed as a separate category because it is not fully a college, and it has some similarities with workplace training, specifically because it provides training on how to perform job-specific tasks. Further, one study (Cattaneo & Boldrini, 2017) took place in a dual vocational educational program and is counted under college or university and workplace. General public includes a study of YouTube videos and a study (Molnar, 2017) that did not identify the study environment, but was not school or work related. Only two studies were situated in the workplace despite the prevalence of the use of instructional video for workplace training. While the ideas learned in one context can inform research in other contexts, there is a need for more research into instructional video in the workplace.
Cognitive Theory of Multimedia Learning (CTML)

Fully half of the studies did not have a theoretical grounding for their research. The other half used theories commonly found in instructional design such as cognitive load theory, Bandura’s theory of observational learning, constructivist theory, and the cognitive theory of multimedia learning (CTML). For this paper, I am limiting the discussion to CTML, which describes how people learn when faced with materials that combine words and images (Mayer, 2021). The words in multimedia instruction can be written or spoken, and the images can be still images or moving as in animations or video. CTML has many different principles that explain how people learn and interact with multimedia. For example, the signaling principle states that “important information should be highlighted through signaling” (Chen & Thomas, 2020, p. 2149). Another example is the dynamic drawing principle which states that “people learn better from a video lecture when the onscreen instructor draws graphics on a board while lecturing rather than referring to already drawn graphic” (Mayer et al., 2020, p. 841).

Since videos are multimedia in format, many researchers investigating instructional video have used the principles of CTML to inform their research. The literature demonstrated that these principles are limited in their application to instructional video, and this section discusses a few such examples.

One example comes from Chen and Thomas (2020) who investigated the principles of temporal contiguity and signaling. According to the temporal contiguity principle, the audio and visual of an action should be presented at the same time. The signaling principle says that “important information should be highlighted through signaling” (Chen & Thomas, 2020, p. 2149). Chen and Thomas compared the conditions of viewing an instructor drawing during a video to viewing partial motion (a series of still images similar to PowerPoint animation) and static images for lecture videos teaching concepts related to drag and lift with airplanes. In all of the videos, the images were the same, but the videos used different levels of animation or drawing. The full-motion condition showed images being drawn while the lecturer talked. The partial motion condition had a video of a PowerPoint slide where the appropriate image appeared as the lecturer spoke. The static image condition showed one static image (i.e., a PowerPoint slide) with all of the images shown from the beginning so that the learner had to know where to look as the instructor spoke. All participants saw three videos with one of the videos in each condition and answered a questionnaire after each video. Videos showing the instructor drawing had an effect on the learner’s level of interest in the video but had no effect on learner’s ability to pass a knowledge test (i.e., recall) or to apply the knowledge to a different situation (i.e., transfer).

Another CTML-related study (Fiorella et al., 2019) looked at the principle of gaze guidance, which says that “people learn better from a video lecture when the onscreen instructor shifts gaze between the audience and board while lecturing rather than looking only at the audience or board” (Mayer et al., 2020, p. 841). Fiorella et al. (2019) found that eye contact with the camera (i.e., students) led to better performance than eye gaze as an attention tool (i.e., gazing on what students should look at). Fiorella et al. investigated the effect of the instructor eye contact in lecture videos about human kidney function. Participants were undergraduates from educational psychology, and there were two conditions: (a) a conventional whiteboard with instructor’s back to the camera and (b) a transparent white board with instructor facing the camera. Participants completed a retention test, transfer test, and a lecture engagement questionnaire. The eye contact group outperformed the traditional whiteboard group on transfer,
but not on retention. The eye contact group also reported higher levels of engagement. There was no difference in self-reported mental effort or perceived difficulty. Fiorella et al.’s findings run counter to the principle of gaze guidance. In addition, Fiorella et al. discussed how the findings indicate that the instructor’s presence on screen may provide benefits beyond helping learners know where to focus their attention.

The Role of the Instructor

Kokoç et al. (2020) investigated the role of the instructor on screen when teaching software and found that the instructor on screen was most helpful to the learners who had the greatest difficulty in focusing on the content. Kokoç et al. investigated different types of video lecture (voice over, picture-in-picture, and screencast) in teaching Python (i.e., software) to undergraduate students. The voice-over condition showed the PowerPoint slides that the instructor used for the lecture, but did not show the instructor’s face. The picture-in-picture showed the PowerPoint slides along with a small image (video) of the instructor talking throughout the lecture. The screencast showed actual usage within Python as the instructor described what he was doing. Prior to beginning the study, participants took a computer-based test to determine their sustained attention levels. The lowest scores overall were from the low attention students in the screencast group. The finding related to the static images is particularly interesting given the dynamic drawing principle of CTML, which states that “people learn better from a video lecture when the onscreen instructor draws graphics on a board while lecturing rather than referring to already drawn graphics” (Mayer et al., 2020, p. 841). Learners with high attention had statistically significantly higher scores with videos that included the instructor’s face along with the screencast as compared with videos that did not include the instructor’s image. Learners with low attention also did better with the videos that had the instructor’s face. According to Kokoç et al. these improved outcomes suggest that the instructor on screen may support the development of a social connection between instructor and student.

Ramlatchan and Watson (2020) also looked at the role of the instructor on screen. They used five different conditions for the same 20-minute video presentation about social media: instructor only, slides only, video switching (alternating view between the instructor and the slides), dual windows (slide and instructor shown side by side), and superimposed (the instructor superimposed in front of the slides). After watching the video, participants answered a questionnaire about the instructor’s credibility and nonverbal immediacy. Instructor credibility is a combination of competence, concern for the learner, and trustworthiness. Nonverbal immediacy is a combination of different nonverbal communication techniques such as hand gestures, facial expressions, and vocal quality. The group that scored the highest for credibility was the dual windows group followed by the superimposed group. The instructor only scored the lowest for credibility. Thus, the slides only condition scored higher for instructor credibility even though the participants could not visually see the instructor. As for nonverbal immediacy, the highest score was the superimposed video followed by video switching, and then the instructor only. Ramlatchan and Watson suggested that having the instructor on screen is not enough--students expect to see slides as part of a lecture. The studies from Kokoç et al. (2020) and Ramlatchan and Watson demonstrate that the issues related to the instructor on screen are complex.
Importance of Culture

One interesting finding from the initial literature was related to the importance of culture in the design of instructional video. *Culture* is a set of beliefs, norms, and practices that are accepted by a group of individuals (Schein, 2017). There are local cultures, national cultures, and organizational cultures, and even occupational subcultures. It might be tempting to avoid considering the culture, especially since only one study (van der Meij, 2019) looked at the impact of culture within the context of instructional video. However, that one study demonstrates the importance of culture in the design of instructional video.

van der Meij (2019) researched the role of national culture on instructional video that teaches software. This is particularly interesting since one might assume that learning software is a culturally neutral activity. van der Meij investigated the use of advance organizers in video, and chose to conduct the study in China with Chinese students because: (a) there was a growing demand for Chinese developers to create training materials and (b) the design of the instruction varied depending upon whether it was created by Chinese or Western designers. According to van der Meij, Western students prefer to learn a single solution at first while Chinese students prefer to be presented with several different solutions. Thus, while a video created for a western audience would present a single approach, a Chinese audience would assume that the presentation of one single approach meant there was only one solution. van der Meij discussed the importance of considering the culture of the learning audience to design instructional videos that align with the learner expectations.

In order to consider the cultural needs, we need to look at what the broader literature on the role of national culture, organizational culture, and occupational subculture on training, regardless of whether the training uses video. I now discuss a few examples from the literature.

**National culture** - Knassmüller and Veit (2016) investigated how national culture affects hiring and promotion practices which then impact training participation and outcomes. Knassmüller and Veit looked at the training participation and outcomes among civil servants in Germany, Austria, Switzerland, and the Netherlands. In Germany and Austria, people perceived training as something that people did when and if they lacked expertise. Thus, training among Germans and Austrians was a sign of weakness or inexperience, and the civil servants in these countries were less likely to attend training.

**Organizational culture** - One aspect of organizational culture that affects training participation and outcomes is the hiring and promotion practices. Rodman et al. (2020) found that highly competitive promotion practices could lead to a culture where knowledge sharing is less likely, even if knowledge sharing is the critical learning modality within the organization. Rodman et al. studied the United States Coast Guard (USCG) where most formal training occurs prior to going to sea (i.e., afloat), and most on-the-job training occurs through knowledge sharing among USCG members when they are afloat. While all members of the afloat community seek to command a ship, a very small number of those who are qualified will be promoted to that level. Rodman et al. found that the competitive nature of promotion practices in the USCG led people to feel they needed to present themselves as being the best at all times, which then made them less willing to share knowledge with others (admitting mistakes could have detrimental effects on the sharer’s career growth). At the same time, all participants mentioned needing shared knowledge more when they were afloat. In other words, the workforce recognized that knowledge sharing was an important component of their professional growth and development, but they did not feel comfortable sharing their knowledge with others.
Occupational subculture – An occupational subculture is a group of professionals who have cultural norms and practices related to the profession. Occupational subcultures have “their own unique sets of values, which influence what should be learnt, when and how, within their own group” (Mak & Hong, 2020, p. 238). These unique sets of values can sometimes conflict with those of the larger organization (Becker, 2018). In one example, the occupational subculture negatively affected firefighters’ willingness to change their professional practice (Lucas & Kline, 2008). Lucas and Kline (2008) conducted a case study understand how occupational subculture among emergency medical services (EMS) professionals and firefighters affected training participation and outcomes. The municipality in the study sought to move from a volunteer fire department that had integrated emergency services to providing emergency services by cross-trained EMS professionals and firefighters. The participants included management, firefighters, and EMS professionals. The firefighters had a hierarchical structure and used this as an excuse to resist change, saying that they wanted to learn to do things differently but they could not because they were beholden to “tradition” (Lucas & Kline, 2008, p. 283). The EMS professionals had a culture that was more flexible where individuals made decisions on their own; as such, it was difficult for the EMS professionals to understand or relate to the challenge facing the firefighters.

The studies from Rodman et al. (2020), Knassmüller and Veit (2016), and Lucas and Kline (2008) do not address the role of culture in the design of instructional video, but they do indicate that culture plays a role in both training participation and outcomes. And if that training is provided in video form, then the design of the instructional video should consider the culture of the learners.

Instructional Video Design Choices

There are many factors that need to be considered when designing an instructional video. In order to understand all of these factors, I created a schematic that presents the different film techniques, instructional strategies, and genres that were discussed in the literature (Figure 1). The film techniques are divided into categories based upon approaches to creating video, and a video might include one item from each of the three subcategories. For example, a video might have the instructor making eye contact while showing the action from a face-to-face point of view, with full motion. Some of the combinations might be more beneficial to learners than others, and other combinations might be difficult to achieve.

The instructional strategies are divided based upon whether the focus of the strategy is for learning, practicing, or retention. And the genres are divided into instructional genres and video genres. As with film techniques, the subcategories may be combined in different ways.

The schematic does not include the instructional aspects of learning objectives, the type of learning (e.g., Gagné et al.’s (2005) five varieties of learning: intellectual skills, verbal information, cognitive strategies, motor skills, attitudes), the needs of the learners, and any unique environmental or cultural factors. For example, if the learning objectives include developing the ability to use a particular software (i.e., intellectual skills and motor skills), then the video should provide support for skill development. The video type might be Screen Cast or Slide Show. The instructional strategy for retention might be to: (a) include segmentation, which is when a video automatically pauses and the learner must take some action to make the video continue, in the video and (b) build testing into the video sequence. The film technique might involve including the instructor on screen. This is just one possibility for an approach to instructional video for software skill development, but there are many other combinations and
permutations that might work equally well. It is possible that certain combinations are best suited to specific cultures and for specific varieties of learning, but that research was not found in this literature review.

Figure 1

*Instructional Video Design Choices*

The diagram does not present desired connections or otherwise indicate the relationships between the different categories and subcategories, even apart from learning objectives and culture. The research is in its infancy with respect to determining these relationships and connections. As such this an initial diagram that will be revised as the field matures.

Design-Based Research

The problem of determining which instructional video approach(es) to use for a particular learning context is daunting if we rely upon research methods that separate out one (or even a few) components in order to determine which combinations provide the best outcomes. Further, conducting such research outside of the real-world context limits the application of the findings to those real-world environments. And instructional video is meant to be used with real learners in a real, and specific context. We need a research approach that is focused on solving complex challenges and addressing problems of practice. Design Based Research (DBR) provides us with a research approach that can help us meet this challenge. But DBR offers something else of significance: the ability to solve a real-world challenge even as we are researching the approach. Design Based Research is sometimes referred to as Education Design Research (EDR), and is a systematic approach to addressing problems of practice by designing and implementing learning
solutions (i.e., interventions) while simultaneously building theory (McKenney & Reeves, 2019; Plomp & Nieveen, 2013).

As Figure 1 indicates, there is not one single problem of practice, but rather many problems of practice. Different contexts and different learning needs require different approaches, and research is needed for all of these. The integrative learning design framework (ILDF) is a DBR approach that allows for addressing local problems of practice through research and then researching those findings in broader contexts. Figure 2 depicts the ILDF (Bannan-Ritland, 2003). The four main phases within the ILDF are: informed exploration, enactment, evaluation: local impact, and evaluation: broader impact. Throughout each phase the researcher works with the community impacted by the problem in a co-design effort. Each phase is iterative and recursive. In addition, the entire process is recursive and iterative, allowing for researchers and practitioners to move forward and backward as needed.

Figure 2
Integrative Learning Design Framework (ILDF)

Note. From “The role of design in research: The integrative learning design framework” by Bannan-Ritland, 2003, Educational Researcher, 32(1), p. 22. Copyright by the Author and Educational Researcher.

During the informed exploration phase, the researcher-practitioner team are building their understanding of the specific training problem, and formulating their initial ideas about what is needed to solve the problem (Bannan-Ritland, 2003). They are identifying the requirements or features that will be needed. This stage includes activities typical of a needs analysis such as stakeholder identification, literature review, analysis of the current environment, site visits (or other approaches to building cultural understanding), and interviews/focus groups.

During the enactment phase, the research-practitioner team is designing and developing a solution (Bannan-Ritland, 2003). This includes developing prototypes and testing out different designs. The designs will include approaches for: delivering the training, collecting data, and building understanding that can be applied to other (broader) environments.

The evaluation: local impact phase is when the researcher is looking data collected during a more in-depth implementation phase (Bannan-Ritland, 2003). The researcher is evaluating the intervention and the theoretical underpinnings that were used to design the intervention. As appropriate, the team may do some additional analysis and exploration tasks. One important distinction between DBR and traditional research approaches is that DBR allows for, and even expects, that there will be changes to the design along the way. If something is not working or if tweaking the design will improve the outcomes, then the design team will likely make changes.
Throughout the entire process, the team is collecting data and keeping records. After all, the purpose of DBR is to develop an educational solution that addresses the problem of practice and can be replicated in other broader contexts. During the local impact phase, it is important to note which aspects of the design are successful in the current environment but might not be workable in other contexts. For example, if the design is highly successful, but there are cultural nuances that help to make this work, then those cultural components need to be noted.

The last phase is when the intervention is implemented and examined in a broader context (Bannan-Ritland, 2003). This is when the researcher shares what was learned in the local context in a manner that enables others to use it. The sharing includes publications and further study. This could mean developing an instructional video solution for one form of workplace training (e.g., ethics) and then testing out those approaches in other forms of training (e.g., sales or marketing). It might involve looking at how an approach in one cultural context can be applied to other cultures, or even taking an approach that worked in one industry and testing it in other industries. The broader impact possibilities are endless.

Conclusion

Instructional video is an important component for training. Yet, the instructional design field has limited understanding of which approaches to video design are best for different learning needs and contexts. Research is needed in practice, and DBR is a research approach that can be used to solve local training needs while building theory and understanding.

References


