Full length article

Instructor presence in instructional video: Effects on visual attention, recall, and perceived learning

Jiahui Wang[*], Pavlo D. Antonenko[**]

University of Florida, School of Teaching and Learning, Gainesville, FL, USA

ARTICLE INFO

ABSTRACT

In an effort to enhance instruction and reach more students, educators design engaging online learning experiences, often in the form of online videos. While many instructional videos feature a picture-inipcture view of instructor, it is not clear how instructor presence influences learners' visual attention and what it contributes to learning and affect. Given this knowledge gap, this study explored the impact of instructor presence on learning, visual attention, and perceived learning in mathematics instructional videos of varying content difficulty. Thirty-six participants each viewed two 10-min-long mathematics videos (easy and difficult topics), with instructor either present or absent. Findings suggest that instructor attracted considerable visual attention, particularly when learners viewed the video on an easy topic. Although no significant difference in learning transfer was found for either topic, participants' recall of information from the video was better for easy topic when instructor was present. Finally, instructor presence positively influenced participants' perceived learning and satisfaction for both topics and led to a lower level of self-reported mental effort for difficult topic.

1. Introduction

With the continued expansion of online learning in K-12 and higher education (Picciano, Seaman, Shea, & Swan, 2012), the ability to support all learners in online learning environments is unprecedentedly important. It has been reported that over 5.8 million students have taken at least one online course in higher education (Allen, Seaman, Poulin, & Straut, 2016, pp. 1–4). However, lack of adequate teacher presence is a common problem in online learning environments (Garrison, 2007). Strategies employed to mitigate this issue include instructor introductions to learning modules, synchronous meetings, virtual office hours, consistent presence in course discussions and prompt instructor feedback. Striving to enhance student engagement and perception of instructor presence in online learning, educators are placing much emphasis on designing and developing online videos that present learning content and frequently integrate the instructor as a picture-in-picture effect within the frame. Some instructional

videos, however, particularly those in the pencast format (Sowa & Thorsen, 2015), do not include an embedded video of the instructor, relying on step-by-step hand writing and voice-over narration by the instructor. A good example of an instructional pencast is the highly popular Khan Academy™ video series, which started out as an online resource offering instructional videos in mathematics and has now expanded to include statistics, chemistry, physics and other academic subjects. One prominent feature of Khan Academy™ is that unlike many other instructional videos, particularly those in the lecture format, it is designed without explicit instructor presence.

Many instructional videos integrate a video of the instructor (e.g., Cousera™, edX™) and this design decision comes at a substantial production cost. Theoretical propositions and empirical evidence for the support of incorporating instructor video in instructional materials are limited and mixed. For instance, the image principle of the Cognitive Theory of Multimedia Learning suggests that people do not learn more deeply when the speaker's image is provided in the instructional presentation (e.g., Mayer & DaPra, 2012; Mayer, Dow, & Mayer, 2003). It should be noted, however, that the image principle was tested using low-embodied or high-embodied animated pedagogical agents, rather than actual instructor videos, so little is currently known about the effects of course instructor presence in instructional videos. What is

* Corresponding author. University of Florida, Norman Hall G518F, 1221 SW 5th Avenue, Gainesville, FL, 32611, USA.
** Corresponding author.
E-mail addresses: jwang01@ufl.edu (J. Wang), p.antonenko@coe.ufl.edu (P.D. Antonenko).

apparent is that students report enhanced engagement when instructional videos include a talking head of the instructor compared to those videos that do not (Guo, Kim, & Rubin, 2014; Kizilcec, Papadopoulos, & Sritanyaratana, 2014).

The current study examined how college students learned with the instructional videos produced by Algebra Nation™, an online community for learning mathematics used by hundreds of thousands of students. The main frame of each video is devoted to a Khan Academy™ style pencast, whereas the bottom right-hand corner always shows a shoulder-up view of the instructor (chosen by the student from a list of about four available instructors of different races and gender). The instructor's shoulder-up video shows the body language and facial expressions of the instructor explaining the content, while the rest of the frame presents a synchronized view of the instructor's hands spelling out and diagramming the problems, concepts, and procedures. The research approach used in this study is novel because in addition to data on learning outcomes (retention and transfer of knowledge), it generated data on the process of learning (visual attention distribution using eye tracking), as well as students' perceptions of their learning with videos on easy and difficult mathematics topics that integrated instructor video and those that did not.

2. Conceptual framework

2.1. Learning and engagement

An important theoretical perspective informing research on the effects of instructor presence in instructional videos is Cognitive Theory of Multimedia Learning (CTML, Mayer, 2014). According to CTML, human memory can be divided into sensory, working and long-term systems. Sensory memory selects and stores relevant visual and verbal information that is received via vision and hearing. Working memory is a central processing unit to process incoming information and integrate it with prior knowledge that has been stored in the long-term memory. Long-term memory stores schemas, or mental structures to organize knowledge. Baddeley's working memory model suggests that working memory has limited capacity (Baddeley & Hitch, 1974) allowing only about four items to be processed at a time (Cowan, 2001). Working memory is also assumed to have sub-units to process different types of information: visuospatial sketchpad for processing visual input, and phonological loop - auditory information. In the context of learning with an instructional video, the narration provided by the instructor would be considered as auditory information to be processed by the phonological loop and information displayed on the screen would constitute visual information processed using the visuospatial sketchpad.

Instructor video embedded in the main instructional video frame is a set of visual stimuli that provide primarily nonverbal communication cues. It is acknowledged that nonverbal communication plays an important role in interpersonal interaction (Argyle, 1988) and facilitates face-to-face mathematics learning (e.g., Alibali & Nathan, 2012). The utility of nonverbal communication also extends to online learning. In the context of instructional video, the image/video of instructor may result in deeper cognitive processing of learning content due to the activation of social interaction schema (Clark & Mayer, 2016). The instructor provides means of nonverbal communication such as mutual gaze, gesturing, and facial expressions. These nonverbal cues could support the cognitive processing of verbal information that is narrated by the instructor, thus improving comprehension. As these nonverbal cues constitute visual information, processed primarily by the visuospatial sketchpad, they should not interfere with the processing of auditory information (e.g., narration), which is handled by the phonological loop. In fact, instructor narration and visual presence in the form of instructor video possibly complement each other as they are processed by different channels (e.g., auditory and visual) and could potentially support information processing in two separate channels resulting in enhanced comprehension of the material. Furthermore, social agency theory suggests that social cues in multimedia presentations lead learners to feel as if they are interacting with another person (Cui, Lockee, & Meng, 2013). From this perspective, social cues in the video replicate the social aspects of human interaction, and this may induce beneficial socio-emotional responses in the learner.

Several studies have examined the influence of instructor presence on learning and perceptions; however, overall, the results appear to be tentative and inconclusive. Evidence of positive effect was provided by Chen and Wu (2015), who used an experimental design and compared the influence of three types of videos on learning: voice over (i.e., instructor's image in the upper left corner of the screen), lecture capture (i.e., a video recording of the lecture) and picture-in-picture. Participants each watched three learning units on document writing presented in each experimental format. Results indicated that performance on recall and transfer of learning with picture-in-picture and lecture capture types was superior to that related to the voice-over type. The three types of video did not cause significantly different levels of positive or negative emotions among participants. On the contrary, Homer, Plass, and Blake (2008) conducted an experiment in which undergraduate students viewed one of two versions of a computer-based multimedia presentation on child development: one included a lecturer with synchronized slides, and the other consisted of slides with audio narration. They compared learning in the two conditions using measures of recall and transfer of knowledge, as well as a social presence questionnaire. No significant difference was found in learning or social presence by including a lecturer in slides with audio narration. Kizilcec et al. (2014) investigated how adding the instructor to instructional video influences undergraduate and graduate students' perceptions and learning on a topic in organizational sociology. Although learners strongly preferred video instruction with instructor presence and perceived it as more educational, they did not perform significantly better on short-term or mid-term recall tests compared to the control condition without instructor presence.

Besides using experimental designs, scholars have also mined Massive Open Online Course (MOOC) server logs and examined the influence of instructor presence in Coursera™ (Bhat, Chinprutthiwong, & Perry, 2015) and edX™ (Guo et al., 2014) MOOC platforms. In a large-scale study of MOOC videos based on 6.9 million video watching sessions across four courses on the edX™ MOOC platform, Guo et al. (2014) examined two proxies for engagement: engagement time (i.e., video watching session length) and problem attempt at follow-up problems. They found students were engaged more with videos that intersperse an instructor's talking head, compared to videos with PowerPoint™ slides alone. Interestingly, the study also suggested that some learners were concerned about the "jarring" effect of having to switch repeatedly between talking head and on-screen text. In a similar study, Bhat et al. (2015) used clickstream data from one Coursera™ course to analyze the engagement (i.e., video watching time, discussion forum visits following a lecture view), motivation (i.e., certificate-earner proportion, fraction of lectures and quizzes that the learner viewed and submitted) and navigational patterns of learners upon being presented with lecture videos incorporating an instructor video in two formats: (a) where the instructor is positioned right next to the slide and seamlessly interacts with the content, and (b) where the instructor appears in a fixed window at the lower left corner of the screen, alongside the content window.