The Instructor's Face in Video Instruction: Evidence From Two Large-Scale Field Studies

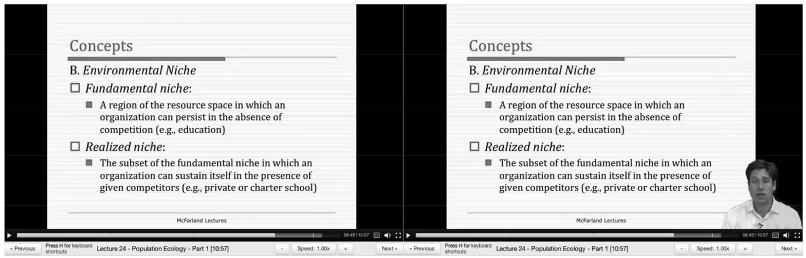
René F. Kizilcec, Jeremy N. Bailenson, and Charles J. Gomez
Stanford University

Multimedia learning research has established several principles for the effective design of audiovisual instruction. The image principle suggests that showing the instructor's face in multimedia instruction does not promote learning, because the potential benefits from inducing social responses are outweighed by the cost of additional cognitive processing. In an 8-week observational field study ($N = 2{,}951$), online learners chose to watch video lectures either with or without the instructor's face. Although learners who saw the face reported having a better lecture experience than those who chose not to see the face, 35% watched videos without the face for self-reported reasons including avoiding distraction. Building on these insights, the authors developed a video presentation style that strategically shows the face to reduce distraction while preserving occasional social cues. A 10-week field experiment ($N = 12{,}468$) compared the constant with the strategic presentation of the face and provided evidence consistent with the image principle. Cognitive load and perceived social presence were higher in the strategic than in the constant condition, but learning outcomes and attrition did not differ. Learners who expressed a verbal learning preference experienced substantially lower attrition and cognitive load with the constant than the strategic presentation. The findings highlight the value of social cues for motivation and caution against one-size-fits-all approaches to instructional design that fail to account for individual differences in multimedia instruction.

Keywords: multimedia learning, video instruction, image principle, cognitive preference, massive open online courses

By the beginning of the 21st century, there were few technological and financial barriers to the production and distribution of instructional videos. The rapid proliferation of instructional video during that time was, therefore, not surprising, in particular with high-fidelity communication media growing increasingly ubiquitous. This trend raises questions about the effective design of audiovisual instruction, especially in light of novel elements of instructional design that have recently become affordable but have received little attention from researchers. Online, asynchronous, computer-mediated communication platforms for educational content, such as massive open online courses (MOOCs), have reached millions of people around the world. One of the largest MOOC providers, Coursera, attracted 2.9 million learners from more than 220 countries in its first year (Waldrop, 2013) and reached seven million learners with over 600 online courses a year later (Koller & Ng, 2014). Many MOOCs rely heavily on audiovisual instruction, further emphasizing the need to study effective designs for multimedia instruction.

The majority of these online courses rely on audiovisual instruction supplemented by regular individual or collaborative homework assignments. The format of audiovisual instruction widely varies across courses, but one popular template is a series of roughly 12-min lecture videos showing lecture slides, overlaid with annotation and the instructor's voice (see Guo, Kim, & Rubin, 2014, for a review of MOOC video properties and their effects on watching behavior). Whereas early lecture videos rarely included an embedded video of the instructor's face, an increasing number of online courses have added a "talking head" in their lecture videos as a picture-in-picture (see Figure 1). This approach has become widely adopted by course producers, such that the majority of recently produced video lectures incorporate a video of the instructor in some form. Although the instructor's face is shown in much of today's audiovisual instruction, the format of presentation varies significantly between courses. Within the confines of practical constraints, such as the space on slides and available recording technology, the presentation of the instructor's face tends to be informed by pedagogical design principles.



By mid-2014, Stanford University Online production manager Wes Choy reported that an instructor's face was present throughout video lectures in 22% of Stanford courses and that the trend in instructional video production was moving toward only showing the face at the beginning and end of lecture videos (personal communication, April 29, 2014). Constantly presenting the instructor's face is thought to enhance learning and motivation in two ways. First, it provides social cues, and, second, it increases comprehension by supplementing the instructor's narration with a nonverbal communication channel that allows for lip reading. However, the constant presentation of the instructor's face can also be distracting and potentially overload the learner with visual information, especially at times when the instructor would like to draw learners' attention to information on the lecture slides. A design strategy that may reap the benefits of social cues and nonverbal communication, and reduces the cognitive burden on learners, would be to show the instructor's face strategically. The presence of his or her face would act as a cognitive aid to direct learners' attention to relevant content during instruction. That is to say, learners would focus on the slides while the face was absent and pay attention to the instructor while the face was present. The presence of the instructor's face would act as a heuristic to guide learners' focus of attention. Hence, the potential benefit of strategically presenting the instructor's face would be to reduce the amount of mental effort required to process lecture materials.

Previous research has investigated the effect of adding a video of the instructor's face in the corner of instructional video. The presence of the face was found to induce higher levels of cognitive load but no significant differences in recall and transfer learning outcomes (Homer, Plass, & Blake, 2008). On further inspection by Homer and colleagues, the effect on cognitive load was moderated by learners' cognitive learning style. For instance, learners with a high visual preference experienced greater cognitive load when the instructor's face was absent, whereas those with a low visual preference experienced greater cognitive load when the instructor's face was present. A recent eye-tracking experiment found that, on average, learners spent 41% of their time looking at the face when it was shown in the video (Kizilcec, Papadopoulos, & Sritanyaratana, 2014). They also dwelled on the face for longer than on the lecture slides but switched frequently—once every 3.7 s—between fixating on the face and on the slides. These behavioral data suggest that the face becomes the primary visual stimulus when shown during instruction and that learners attend to slides to retrieve corresponding information. Despite its substantial effect on watching behavior, Kizilcec et al. (2014) did not find an effect of presenting the face on the ability to recall information. Yet, according to survey responses, participants in the study strongly preferred instruction with the face and thought it was more educational than instruction without a face present.

This article extends the existing literature on the image principle (Mayer, 2001), which is reviewed in the next section, by examining the behavior and attitudes of a large, diverse sample of adult learners over two multiweek periods in response to different presentation styles. The first study was designed to uncover which presentation style learners prefer in a realistic learning scenario and the rationale behind their preference. Learners chose to watch lecture videos either with or without the instructor's face after balanced exposure to both presentation styles. In the second study, we examined the outcomes from the first study, in particular the finding that a substantial number of learners preferred watching the video without the face, because they found the instructor's face distracting. Because the majority of learners still preferred videos with the instructor's face, we developed a strategic presentation style that omits the instructor's face when it is potentially most distracting. Learners in a 10-week online course were randomly assigned to lecture videos with either a constant or a strategic presentation of the instructor's face. Learning outcomes, behavioral data, and self-report data, including cognitive load and perceived social presence, were collected for each learner. These studies represent unique contributions to the literature in that, previously, there had been no longitudinal investigations of learners' attitudinal and cognitive responses to and choice of multimedia presentation styles in real-world learning environments.

First, we review relevant theoretical and empirical work in educational psychology and related disciplines. Then, present the first study, which was exploratory in nature and yielded insights that informed the design and research questions of the second study. The presentation of the second study is prefaced by a theory-driven development of several research questions on the comparison between the constant and the strategic presentation styles. A final discussion synthesizes the key findings and implications from both studies.

Theory

The ultimate goal of instruction, whether in person or digitally mediated, is *learning*, defined from a learning science perspective as a change in a person's knowledge caused by an experience (Bransford, Brown, & Cocking, 2000). There is, however, no definitive conceptualization of the process of learning—the mechanism by which a learning experience translates into a change in knowledge. According to the traditional view of schooling, known as *instructionism*, learning is the process of knowledge acquisition in which knowledge is a collection of facts and procedures and the learner is a passive recipient of information during instruction (Papert, 1993). In contrast, following the knowledge construction conceptualization of learning, learning is building mental representations by actively engaging in cognitive processing, such as directing attention to relevant information and mentally organizing and integrating new information with prior knowledge (Bransford et al., 2000).

Theories of multimedia learning generally assume that learning is an active process of knowledge construction in which the goal of instruction is to guide learners' cognitive processes (Clark & Mayer, 2011). To optimize for cognitive processing, a model of how cognitive processes unfold in the learner's brain is needed. Mayer (2005) proposed a model, known as the *cognitive theory of multimedia learning*, which relies on three assumptions grounded in evidence from cognitive science. First, visual and auditory information are processed in separate channels (dual channels; see Baddeley's, 1986, 2003, model of working memory). Second, only a limited amount of information can be processed in each channel at any time (limited capacity; Miller, 1956). Third, learning occurs when a learner is engaged in certain cognitive processes (active processing)—namely, attending to certain verbal and auditory information with his or her senses—then organizing the selected information into coherent mental representations and, finally, integrating mental representations from the two channels with prior knowledge. On the basis of this model of cognitive processing, empirical evidence of how instruction can optimize cognitive processes has been synthesized into principles of multimedia instruction (see Clark & Mayer, 2011).

Cognitive Load Theory

Multimedia learning theory has been influenced by cognitive load theory (Sweller, 1988; Sweller, van MerriËnboer, & Paas, 1998), an educational theory that provides a conceptual model for the kind of cognitive processes that occur during instruction. This is useful for understanding the potential effects of features in multimedia instruction on cognitive processing. Cognitive load is theorized to consist of three additive, interrelated components: (a) intrinsic, (b) extraneous, and (c) germane cognitive load. Learning may only occur when the sum of these three components does not exceed an individual's working memory resources (i.e., cognitive overload). The presentation of the instructor's face can be characterized as inducing germane (enhancing learning) or extraneous cognitive load (hindering learning).

In this article, we assume a model of cognitive load in which (motivated) learners divide all of their available cognitive resources to process the instruction. Accordingly, removing a content-irrelevant element of instruction would redistribute cognitive resources from extraneous to germane processing. Learners, thereby, would gain additional resources for germane processing, and reaching a state of cognitive overload becomes less likely.

Intrinsic cognitive load is a function of *element interactivity* and *learner expertise*, where element interactivity is defined by the required amount of information units held in working memory to comprehend the material. With greater learner expertise, the number of information units that a learner needs to hold in working memory to understand the material is reduced (van MerriËnboer, Kester, & Paas, 2006). For example, teaching arithmetic is expected to induce less intrinsic cognitive load than teaching calculus, but learning arithmetic in second grade and learning calculus in 11th grade may induce similar levels of intrinsic cognitive load. Extraneous cognitive load is not necessary for learning and is typically an artifact of suboptimal instructional design. For example, playing background music during instruction is a likely source of extraneous cognitive load. Germane cognitive load, though not necessary for learning to occur, actively enhances learning (i.e., aids the construction of cognitive structures) and typically results from aspects of instructional design. For example, an instructor's hand gestures to demonstrate forces acting on an object in a physics class can increase learners' comprehension, with the additional visual processing of hand movements causing germane cognitive load.

The cognitive load framework can guide comparisons of different presentation styles of the instructor's face in video lectures. There are arguments for why each style would be expected to induce extraneous or germane cognitive load, which are discussed in a later section.

Relevant Principles for Multimedia Instruction

The personalization principle, which is aimed at motivating learners to exert effort to learn, suggests that adding social cues to instruction can promote learning by fostering generative processing (Clark & Mayer, 2011). The rationale behind this principle is that social cues induce learners to treat the technologically mediated learning experience as if it were human-to-human and, thus, respond socially (see Reeves & Nass's, 1996, computers-as-social-actors paradigm). Because face-to-face communication has historically been the primary mode of knowledge transfer among humans, seeing a face during instruction is a familiar experience that naturally elicits learned social responses. Politeness, for instance, is a social response that would increase active cognitive processing as learners pay more attention to being polite, which leads to improved learning outcomes. Numerous studies provide evidence for this positive effect on learning, for example, as a result of changing text or narration from formal to conversational style (Kartal, 2010; Moreno & Mayer, 2000), using a human voice rather than a computer-generated voice (Atkinson, Mayer, & Merrill, 2005; Mayer, Sobko, & Mautone, 2003), and using polite speech (McLaren, DeLeeuw, & Mayer, 2011; Wang et al., 2008). Considering social cues more broadly, a large body of literature on the role of social presence in online courses highlights the benefits of social presence in terms of perceived learning and satisfaction (Gunawardena & Zittle, 1997; Lyons, Reysen, & Pierce, 2012; J. C. Richardson & Swan, 2003; Tu & McIsaac, 2002).

In light of the empirical evidence for the benefit of embedding social cues in voice, adding a visual social cue in the form of an image of the instructor is a natural extension (see Yee, Bailenson, & Rickertsen, 2007, for a review of research on faces in interfaces). In the image principle, Mayer (2001) summarizes evidence from several studies investigating the effect of adding a realistic or cartoon-like pedagogical agent in multimedia instruction (e.g., Mayer, Dow, & Mayer, 2003; Moreno, Mayer, Spires, & Lester, 2001), which demonstrate mixed results on the effect on learning. The image principle suggests that the potential benefit from inducing social responses by adding an image of the instructor's face tends to be offset by extraneous visual processing that the image inevitably induces, because it does not contain pedagogically relevant information. The coherence principle (Clark & Mayer, 2011) suggests that adding extraneous graphics, such as extraneous video clips (Mayer, Heiser, & Lonn, 2001), hinders learning. If applied to the image of the instructor, the coherence principle recommends removing it unless the benefit from social cues can justify drawing on additional cognitive resources.

Besides inducing social responses, an onscreen agent or video of the instructor may promote learning by directing the learner's visual attention to relevant parts of the screen by, for instance, pointing at a graphic. Highlighting relevant material to direct visual attention is a technique called *signaling* and traditionally leverages typography, color, and the insertion of arrows (Mautone & Mayer, 2001; Mayer & Moreno, 2003). However, even if the pedagogical agent, or the instructor in an embedded video, points at relevant material, the problem of split visual attention remains: The learner is confronted with multiple visual stimuli to process simultaneously, some of which are intimately linked to auditory stimuli (e.g., in the case of the instructor's face and voice).

In summary, this preliminary review of relevant theory and empirical work suggests a tradeoff between the costs and benefits of presenting the instructor's face in video instruction. On the one hand, the face could induce beneficial social responses and support learners' understanding by providing nonverbal cues. On the other hand, the face is an additional visual stimulus that might unnecessarily burden and even actively distract learners. The following study extends prior work by investigating learners' choices and rationales for choosing instruction with or without the instructor's face over time in an online learning environment.

Study 1: Longitudinal Observational Study

The goal of this observational study was to shed light on learners' informed choices of presentation style and the reasoning underlying their choices. To this end, we examined learners' behavioral patterns of engagement with video lectures in a MOOC and their open-response feedback on why they preferred their chosen presentation style. To ensure that learners' choice of presentation style would be a deliberate and informed choice, they were exposed to both presentation styles before making a choice. In the absence of a control group with a randomly assigned presentation style, this study could not provide insights into the effect of choice over presentation styles for learners. Although the psychological process of choosing may itself induce learner effects, the goal of this study was to better understand which presentation styles learners choose in a realistic learning environment over time.

We predicted that a majority of learners would choose the presentation with the face, on the basis of prior work indicating more positive attitudes toward this presentation style relative to instruction without a face (Kizilcec et al., 2014). Although previous work has found higher cognitive load in the presence of the instructor's face (Homer et al., 2008), this finding might or might not be replicated in the presence of self-selection (as opposed to random assignment to conditions). We therefore had no prediction as to whether learners' self-reported affect, effort, and perceived learning would also be higher with the instructor's face than without it when learners self-select into presentation styles. Although observational evidence seldom warrants causal claims, if previously found effects on affect, effort, and perceived learning were not replicated with self-selection, it could suggest that learners' preferences and subsequent choices are partly determined by these constructs. We expected that learners' self-expressed rationales for choosing between presentation styles would provide a richer insight into the thought process around the decision to watch videos with or without the face.

Method

Participants. Participants were enrolled in a MOOC on a topic in sociology on the Coursera platform (http://www.coursera.org/). The course started in September 2013 and lasted for 10 weeks (in total, 19,088 learners enrolled and watched at least one lecture). For the behavioral analysis of video watching in this study, we only considered learners who watched at least 3 min of each lecture video in the 1st week of the course and thereby were exposed to videos with and without the face ($N = 2{,}951$). On the basis of an optional survey (67% response rate), learners were balanced in gender (48% female, 52% male), 37.6 years old on average (1st quartile = 28, $Mdn = 35$, 3rd quartile = 45), and three-quarters of them had 4–6 years of college/university education.

Materials. The basic course material included video lectures (about six videos each week, each video lasting between eight and 20 min), weekly multiple-choice questions, and a final exam that consisted of 100 multiple-choice questions. Moreover, for the purposes of this observational study, the standard course material was extended to include each lecture video in two clearly labeled versions—with and without the instructor's face—starting in the second week of the course. A video recording of the instructor's face was continuously present in the lower right corner of the video in one version (i.e., video with talking head), whereas the face was completely removed, leaving only lecture slides with narration in the other version (see Figure 1). The study ran in the first eight weeks of the 10-week course.

Because the design space for the presentation of video instruction is very large, our choice of placing the talking head in the lower right corner requires some justification. Our design choice followed the convention set by the vast majority of lecture videos produced for MOOCs at the time, which presented the instructor's face as a picture-in-picture in the lower right corner. This convention may have originated because it is least likely to obstruct materials on the lecture slide given left-to-right and top-to-bottom writing conventions.

In the 1st week of the course, all learners watched the same videos, with the presence of the face alternating between consecutive videos to expose them to both versions and enable them to make an informed choice in later weeks. In addition, the instructor informed learners in the first video of the course that they would be given a choice of presentation styles as part of a research project to investigate online learners' preferences for video presentation.

The order in which the two versions of each video were presented was alternated, and the two versions of each video were presented in pairs to reduce order effects. For instance, the first and second video and the third and fourth video in the 2nd week were identical except that the instructor's face was shown in the second and fourth video but not in the other two. Each video was clearly labeled as *with face* and *without face*. To check for order effects, the order of presentation was reversed in Weeks 5 and 7, such that the version with the face was presented first in each pair in these weeks.

Measures. A number of behavioral and self-report measures were collected for different samples of learners. Video-watching behavior was observed for learners who watched the 1st week's video lectures in the course ($N = 2{,}951$). Self-report measures were optional and collected for any learner in the course who completed them, including those who did not watch all videos in the 1st week. Response rates varied between 10% and 18%, which is relatively high in the context of open online courses, which are notorious for their high attrition rates.

Video watching. All interactions with video lectures in the course were recorded in a course database with a timestamp. This provided information on which lecture version (with or without the face) was watched. Learners were considered to have watched a video if they spent at least 3 min on the page where the video was shown.

Affect, effort, and perceived learning. In Weeks 2–8 of the course, learners were prompted to complete an optional three-item survey at the end of one video lecture (same lecture with and without face) each week to rate perceived learning, effort, and affect (18% response rate). Learners rated "how much effort it [took them] to watch this lecture" on a labeled five-point scale from *none at all* to *a great deal*; "how much [they learned] from this lecture" on a labeled five-point scale from *nothing at all* to *a great deal*; and "how good or bad [their] experience [was] watching this lecture" on a labeled seven-point scale from *extremely bad* to *extremely good*. Affect, effort, and perceived learning was measured with one item per construct on five- or seven-point scales, such that measures ranged from 1 to 5 or 1 to 7. The effort scale was adapted from the mental effort scale used in prior cognitive load research (e.g., Paas, 1992; Paas & van MerriËnboer, 1994, who measured mental effort induced by different computer-based training strategies; see Brünken, Plass, & Leutner, 2003, for a theoretical account of cognitive load measurement in multimedia learning).

Preference and rationale. At the end of the course, learners were prompted to complete an optional survey, which included three questions about lecture presentation styles (10% response rate). There was a multiple-choice question on "which lecture version [they preferred]" with the following answer options: "with face' lectures," "without face' lectures," "I like having both versions to switch back and forth," and "I don't have a preference." Two open-response questions asked learners "how [they] decide[d] which version to watch" or why [they] deliberately switch[ed] between versions" and "from [their] experience, what [were] the merits and costs of showing the instructor's face compared to not showing it?"

Procedure. Learners enrolled in the course and were informed about the observational study and exposed equally to both presentation styles in the 1st week. Video lectures were released each week in the course, including both a "with face" and a "without face" version in Weeks 2–8.

Statistical considerations. Learners repeatedly reported perceived learning, effort, and affect in different videos throughout the course, and observations were therefore not independent. We constructed a linear model with a fixed effect for whether the face was present or absent in the video and a fixed effect for each video to account for differences across videos other than the presence of the face. The linear model was bootstrapped with 1,000 replications (clustered on learners) to estimate standard errors around the relative increase and establish the null distribution for a z test. Note that Cohen's d is not reported because it is a normalized effect size that depends on the estimated variance, which could have been computed in a number of ways in this study design with repeated measurement.

Results

Although the majority of learners tended to watch videos with the face (57%), a substantial number tended to watch videos without the face (35%), with the remaining 8% of learners watching both versions. The first time learners were given a choice of presentation styles, in Week 2, the distribution across versions was even more balanced: 35% watched with the face, 44% watched without the face, and 21% watched both versions. There was, however, evidence of order effects in the 2 weeks when videos were presented in reverse order (the foregoing percentages were observed in weeks when "without face" was shown first). A larger percentage chose the presentation style that was shown first, such that when the "with face" version was shown first, 70% chose videos with the face, 23% chose without, and 7% watched both. The presence of nonnegligible order effects may reflect some learners' indifference to presentation styles. All percentages were significantly independent, $\chi^2(2, N = 2{,}951) > 180$, $p < .001$.

In addition to observing which videos they opted to watch, learners rated their level of perceived learning, effort, and affect immediately following lecture videos. Summary statistics, statistical tests, and relative effect sizes suggest that learners who watched videos with the face thought they learned more and needed to exert less effort and rated the experience better than did learners who watched videos without the face (see Table 1). In this observational study, learners self-selected into watching videos with or without the instructor's face and as to whether they completed optional surveys. A comparison of self-report measures for learners who watched videos with the face and learners who watched videos without the face was expected to yield biased estimates of causal effects because of to self-selection. Yet the observed differences were significant and nonignorable in effect size, suggesting that those who opted for videos with the instructor's face were, on average, subjectively better off than those who opted to watch videos without the instructor's face.

A subset of learners reported their preferred presentation style and the rationale for their preference in an optional feedback survey toward the end of the online course. Consistent with the overall choice behavior, most learners reported a preference for videos with the face (59%) over videos without the face (16%); 17% liked having both versions to switch back and forth between, and 8% had no preference, $\chi^2(3, N = 2{,}231) = 1{,}892$, $p < .001$. Learners' self-reported rationales for choosing a presentation style were insightful and closely related to the theoretical considerations outlined earlier—though the samples for behavioral and self-reported results were different, and response options did not directly map onto video choices. Some learners expressed the view that the instructor's face only added extraneous cognitive load; one learner explained that "the instructor's face might personalize the content, but since most of what was being described was available in the slides, I did not find it of much value." Moreover, several learners wrote that "[they] started [watching videos] with [the] face but found [they] could concentrate more [without the] face." In contrast, other learners found that "making eye contact with the speaker helped [them] focus" and that it also helped them "feel connected to the instructor as a person." Some learners hinted at the value of lip reading and nonverbal cues in general, explaining that they chose the version with the face, "because for a nonnative speaker, nonverbal communication is easier."

Discussion

The goal of this exploratory study was to shed light on learners' choices of presentation style and the reasoning underlying their choices. Although the majority of learners preferred lectures with the face, one-quarter to one-third of learners chose to watch lectures without the instructor's face. Around 8% of learners actively used both versions. Those who watched videos with the instructor's face reported liking the lectures better, needing to exert less effort, and learning more than those who watched videos without the instructor's face. This replicates in the field what has previously been found in a controlled laboratory environment for perceived learning and affect (Kizilcec et al., 2014), but it is inconsistent with previous work that found higher cognitive load with the face in a laboratory experiment (Homer et al., 2008). This inconsistency may be a result of self-selection bias or differences in the measurement of mental workload. Learners' self-reported reasons for their choices of presentation style highlight the value of social and other nonverbal cues from showing the instructor's face, but they also raise the issue of distraction and call the face's educational value into question.

Many learners reported that following the lecture and paying attention was easier with the face, which is consistent with some neurological and evolutionary accounts that further conceptualize the effect of showing the instructor's face in video instruction. Specifically, the combination of hearing speech and seeing the instructor's face or gesture and speech was found to alleviate learners' cognitive load (e.g., Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001).

The key insight gained from this study is that although the majority of learners preferred watching lectures with the face, there also was a substantial number who preferred not seeing the face. Some learners switched between presentation styles, potentially to balance the benefits and costs of seeing the instructor's face during instruction. This study did not provide insights into the difference between assigning presentation styles and allowing learners to make a choice. Disentangling the effects of these processes could be the subject of future research. Yet the findings suggest that assigning the same presentation style to all learners would be in conflict with the revealed preference of at least one-third of the learners in our sample. These results motivated the next study, in which learners were randomly assigned to presentation styles. We conducted a study to compare the effects of constantly showing the instructor's face with those of strategically showing the instructor's face—a presentation style that attempts to reduce the level of distraction while preserving the social and nonverbal benefits of the instructor's face.

Study 2: Longitudinal Field Experiment

In this study, we tested a technique of instructional design that is intended to reap the benefits of inducing social responses and adding nonverbal cues while reducing the amount of extraneous processing, in alignment with the coherence principle. We investigated the effectiveness of this new approach over time relative to the standard approach of constantly showing the video of the instructor; this was the first longitudinal investigation on the effectiveness of pedagogical agents of its kind.

Research Questions

Building on the theoretical framework presented earlier, we posed research questions regarding the effects of constantly and strategically presenting the instructor's face on cognitive load, social presence, learning outcomes, and attrition.

Cognitive load. Per the coherence principle (Clark & Mayer, 2011; Mayer et al., 2001), the strategic presentation, relative to the constant presentation, would be expected to free up a larger amount of learners' limited cognitive resources for processing content-relevant information. This is because the instructor's face is less present, which should be reflected in lower levels of reported cognitive load. Moreover, the signaling principle (Mautone & Mayer, 2001; Mayer, Dow & Mayer, 2003) suggests that directing learners' attention to critical elements of the lecture—typically by highlighting text, drawing arrows, or making emphases in narration—reduces the necessary amount of extraneous cognitive processing by serving as a cognitive aid to the learner. From the classical conditioning perspective, the instructor's intermittently disappearing face can be considered a conditioned stim-

Table 1
Means and Standard Deviations (in Parentheses) for Statistical Tests and Relative Effect Sizes With 95% Confidence Intervals for Self-Reported Learning, Affect, and Effort in Video Lectures With and Without the Instructor's Face

Variable	Perceived learning	Affect	Effort
With face	3.52 (0.79)	5.30 (1.00)	2.85 (0.94)
Without face	3.46 (0.82)	5.15 (1.07)	2.93 (0.93)
Test statistic (z)	3.6	6.6	−4.0
p	$<.001$	$<.001$	$<.001$
Effect size	1.7% [0.8%, 2.7%]	2.8% [2.0%, 3.7%]	−2.9% [−1.5%, −4.3%]

Note. Data are based on 21,576 responses from 3,494 learners. Effect sizes and statistical tests were derived from a linear regression with a fixed effect for each video in the course using a learner-clustered bootstrap.

ulus that becomes associated with the unconditioned stimulus of paying attention to the slides, thereby guiding learners' visual attention (Dawson, Schell, Beers, & Kelly, 1982). The strategic presence of the instructor's face is designed to signal to the learner when to direct attention to the lecture slides. Therefore, this was expected to reduce cognitive processing relative to the constant presentation and should also have resulted in lower levels of cognitive load. Nevertheless, signaling with arrows or by circling key phrases is a more natural, direct signaling cue than the presence or absence of the instructor's face, which might take learners' more time to identify as a cognitive aid. Repeated switching between the presence and absence of the instructor's face may incur a cognitive switching cost similar to the extraneous processing incurred by switching between words and pictures, as identified in the spatial contiguity principle (Clark & Mayer, 2011; Ginns, 2006). Research on the effects of such mixing costs suggests that they diminish with enough practice (Strobach, Liepelt, Schubert, & Kiesel, 2012), though it is unclear how these findings translate to the present stimulus in the context of video instruction. All in all, prior work is inconclusive on the likely effects on cognitive load of the strategic relative to the constant presentation. Hence, we posed the following as a research question:

> *Research Question 1:* Is cognitive load higher in the strategic or the constant condition?

Social presence. Similar to a news anchor directly speaking into the camera and, apparently, *at* the viewer, the instructor directly addresses the learner audience, which tends to experience this as a parasocial interaction (Horton & Wohl, 1956). Learners ignore the mediated nature of the lecture experience and perceive the instructor as a social actor, and they respond by mindlessly applying learned social rules (Nass & Moon, 2000).

The personalization principle (Clark & Mayer, 2011) and, more specifically, the image principle (Mayer, 2001), suggest that inserting cues or other design elements that emphasize the presence of the instructor should induce learners to respond more socially by, for instance, inducing more politeness, which effectively leads learners to be more attentive (McLaren et al., 2011; Wang et al., 2008). Given that the social cue of the instructor's face is absent for some periods in the strategic condition, the learner's experience of social presence from the instructor may be weakened relative to that in the constant condition. However, there could be a ceiling effect for the extent to which showing the instructor's face increases feelings of social presence. Thus, the levels of social presence would be similar across presentation styles, because the intermittent presence of the instructor's face in the strategic condition already induces the maximum possible amount. Moreover, the repeated appearance and disappearance of the instructor's face in the strategic condition is likely to draw greater attention to the instructor. This would increase the salience of the instructor's presence to the learner and, hence, increase feelings of social presence. There are, again, competing hypotheses about the effect of the strategic relative to the constant presentation on social presence. We posed the following research question:

> *Research Question 2:* Is social presence higher in the strategic or in the constant condition?

Learning and assessment taking. According to the coherence principle (Clark & Mayer, 2011; Mayer et al., 2001), the strategic presentation would be expected to induce lower levels of extraneous load than the constant presentation, because the extraneous visual cue is present for shorter durations. If a smaller percentage of learners' limited cognitive capacity is spent on extraneous processing, more resources are available for organizing and integrating learning materials. Thus, learners can perform better on recall and transfer tests. From a speech perception standpoint, however, the continuously present face of the instructor allows learners to speech-read by observing the instructor's articulation, facial and manual gestures, and audition. The bimodal nature of human speech perception was famously demonstrated in the *McGurk effect* (McGurk & MacDonald, 1976). The benefit of vision to complement speech perception in a noisy environment is especially large for those who can lip read (MacLeod & Summerfield, 1987). The positive impact of audiovisual integration was also found for the perception of natural language (Dekle, Fowler, & Funnell, 1992). Hence, the instructor's face would be expected to enhance learners' understanding of the lecture, especially when the lecture was taken in in a noisier environment and by nonnative speakers (Hapeshi & Jones, 1992).

Research on social presence (Short, Williams, & Christie, 1976) in online learning (see Cui, Lockee, & Meng, 2012, for a review) has found several positive outcomes associated with increased perceived social presence, such as increased learner satisfaction and perceived learning (Gunawardena et al., 1997; Lyons et al., 2012; J. C. Richardson & Swan, 2003; Tu & McIsaac, 2002). More immersive learning experiences were also found to enhance learning compared with traditional instruction (Dede, 2009). Taken together, we have here another set of conflicting hypotheses about the effect of the presentation of the instructor's face on learning outcomes, and we posed the following research question:

> *Research Question 3.1:* Will learning outcomes be higher in the strategic or in the constant condition?

In an open learning environment such as a MOOC, learners may or may not be motivated to take assessments. The personalization principle (Clark & Mayer, 2011) suggests that additional social cues from the instructor should motivate learners to be more engaged. It is unclear, however, whether showing the instructor's face strategically or constantly in video lectures affects learners' motivation to take assessments. Accordingly, we posed the following research question:

> *Research Question 3.2:* Are learners more likely to take assessments in the strategic or in the constant condition?

Attrition. Attrition in online learning environments is a complex process to capture, especially in MOOCs, where the absence of financial and institutional barriers to enter and exit courses encourages participation at varying levels of commitment (Kizilcec, Piech, & Schneider, 2013). Identifying factors associated with attrition has been the focus of a substantial amount of research (e.g., Packham, Jones, Miller, & Thomas, 2004; Rovai, 2003; Willging & Johnson, 2004). A general finding in the majority of investigations is that attrition is influenced by learners' satisfaction with the learning experience. Factors such as overly high cognitive load or low social presence are likely to reduce a learners' satisfaction with their learning experience and increase their likelihood of leaving a course. Building on the logic behind the effects of the strategic relative to the constant presentation of the instructor's face on cognitive load and social presence, attrition may be higher or lower depending on the direction of the effect on cognitive load and social presence. Therefore, we posed the following research question:

> *Research Question 4:* Is attrition lower in the strategic or in the constant condition?

Learning preference. A number of investigations into multimedia learning have uncovered effects moderated by individual differences, such as expertise (Kalyuga, Ayres, Chandler, & Sweller, 2003). Mayer and Massa (2003) reviewed 14 cognitive measures and found them to cluster into three categories: (a) cognitive ability, (b) cognitive style, and (c) learning preference. Homer and colleagues (2008) found cognitive style to moderate the effect of showing the face of a speaker delivering a talk on cognitive load, such that those with a visual cognitive style experienced lower cognitive load than those with a verbal cognitive style. In a similar manner, when the face was absent, those with a verbal cognitive style experienced lower cognitive load than those with a visual cognitive style. The visualizer–verbalizer hypothesis states that some people are better at processing pictures, whereas others are better at processing words (Mayer & Massa, 2003). However, recent work suggests that a more complex characterization is needed—one based on evidence for two types of visualizers with varying spatial ability (Kozhevnikov, Hegarty, & Mayer, 2002; Kozhevnikov, Kosslyn, & Shephard, 2005).

For the present study, we measured individual differences in learning preference using a single question to assess whether learners preferred receiving new information visually (e.g., in graphics) or verbally (e.g., in written or spoken text). Building on the findings of Homer et al. (2008), we posed the following research question:

> *Research Question 5:* Do individual differences in learning preference moderate the effect of the strategic presentation relative to the constant presentation of the face on (a) cognitive load, (b) social presence, (c) learning outcomes and assessment taking, and (d) attrition?

Method

Participants. Participants were enrolled in a MOOC on a topic in sociology offered on the Coursera platform. The course started in September 2012 and lasted for 10 weeks. A total of 44,432 learners enrolled in the course, 27,855 entered the course, and 12,468 watched at least one lecture video (this set of learners is the overall sample in this experiment); 3,697 learners were active in the last week of the course, and 1,580 certificates of accomplishment were issued to learners who attained a course grade of at least 70%. On the basis of an optional survey (11% response rate), learners were balanced in gender (51% female, 49% male), 63% were between 25 and 44 years of age (19% under 25, 18% over 44), and over three-quarters of them held a bachelor's or master's degree.

Materials. The basic course material included video lectures (about six videos each week, each video was between 8 and 20 min long), weekly multiple-choice questions that were embedded in video lectures and accessible as stand-alone quizzes, and a final exam with over 100 multiple-choice questions.

The presentation of the instructor's face was manipulated in a total of 57 lecture videos to produce 57 lecture videos for the constant face condition and 57 for the strategic face condition. To create the videos, the face of the instructor was recorded (with a blue screen to produce a transparent background) when the instructor recorded the spoken text for the lecture videos (see Figure 1). The instructor's face was present for 33% to 67% of the video in the strategic condition. The decision mechanism underlying the presence and absence of the instructor's face was based on the cognitive aid strategy discussed earlier and is illustrate in Figure 2.

The process of creating the strategic face version was systematic in the sense that all 57 video lectures were individually reviewed and edited on the basis of a predefined rule for when to remove the face and when to keep it. The process, however, was subjective, because all videos were coded and edited by only one researcher.

Measures. A number of behavioral and self-report measures were collected for the sample of learners who were exposed to the manipulation by watching at least one video lecture in the course ($N = 12{,}468$). Response rates for optional self-report measures are reported. Response rates varied between 11% and 21%, which is relatively high in the context of open online courses, which are notorious for their high attrition rates.

Attrition. Attrition was based on the start and stop times of watching video lectures in the course. The start and stop times for each learner were determined using the 5th and 95th percentile of their individual lecture watching times. All lecture-watching events were automatically recorded by the learning environment and stored in a database.

Course grade and assessment taking. The overall course grade was based on multiple sources of assessment, including the final exam; weekly quizzes; peer-graded essays; and, to a small extent, regular forum participation. The final exam consisted of 104 multiple-choice questions covering material from the whole course.[1] The course grade, which could range from 0 to 100, therefore provided a holistic measure of recall and transfer learning for those who were motivated to take assessments (70 was the threshold to receive a certificate of completion). Eighty-nine percent of learners received a zero course grade because they did not complete any assessments in the course, despite engaging in lecture watching. Therefore, we separately report the percentage of learners who completed assessments (i.e., those who achieved a nonzero course grade) and the average course grade for all learners.[2]

Cognitive load. Cognitive load was self-reported twice a week immediately following a lecture video (21% response rate). We used an adapted subset of the NASA TLX measure (Hart & Staveland, 1988) consisting of two items: (a) "How much mental demand did you experience watching this lecture?" and (b) "How much effort did it take you to watch this lecture?" The response scales were 101-point sliding scales with a default midpoint at 50.

[1] Final exam items can be accessed at http://kizilcec.com/rsc/jep2015face.pdf

[2] We intentionally do not condition average grades on whether learners took assessments, because this quantity is subject to selection bias. The conditional-on-positive effect does not have a causal interpretation, even in a randomized controlled experiment (Angrist & Pischke, 2008).

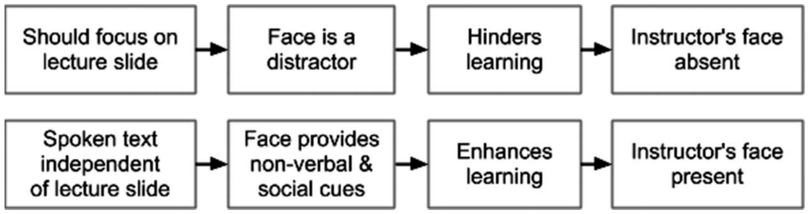


Figure 2. Decision mechanism for the presence and absence of the face in the strategic face condition.

Cognitive load scores were computed by averaging the two items and then transforming them to vary between 0 and 1. Scores were transformed to range from 0 to 1 by subtracting the minimum and dividing by the maximum value. The questionnaire appeared in the same window as the lecture video as soon as the video ended.

Social presence. Social presence was self-reported once each week immediately following the lecture video using a five-item questionnaire (Cronbach's $\alpha = .72$) adapted from several existing social presence measures (17% response rate). Each of the following statements was rated on a five-point labeled scale from *not at all* to *very strongly*: (a) "I felt like the instructor was in the same room as me," (b) "I felt that the instructor was very detached in his interactions with me" (reversed), (c) "I felt that the instructor was aware of my presence," (d) "I felt that the instructor was present," and (e) "I felt that the instructor remained focused on me throughout our interaction." Social presence scores were computed by averaging the five items and then transforming them to vary between 0 and 1. The questionnaire appeared in the same window as the lecture video as soon as the video ended.

Learning preference. Cognitive preference for visual or verbal information was measured with one item from the Learning Styles Index (Felder & Spurlin, 2005): "I prefer to get new information in a. pictures, diagrams, graphs, or maps, or b. written directions or verbal information." Learners were prompted to complete an optional survey at the beginning of the course (11% response rate); 430 reported a preference for verbal information, and 899 reported a preference for visual information.

Procedure. Learners were randomly assigned into the two video conditions at the time of their enrollment on the basis of their user ID after providing consent to participate in research conducted during the course. Out of the learners who watched at least one video lecture, 6,335 were assigned to the constant face condition, and 6,133 were assigned to the strategic face condition. All learners interacted with the same learning platform but were provided with different sets of videos. In the constant face condition, participants viewed video lectures with a constant talking head of the instructor, whereas the talking head was intermittently present on the basis of the strategy described earlier in the strategic face condition.

Statistical considerations. Randomized controlled experiments are the predominant scientific method for discovering causal relationships. Random assignment to conditions creates independence between the experimental manipulation and any covariates, such that the comparison of experimental groups can yield an unbiased estimate of the average causal effect. In a longitudinal field experiment like this, a comparison between experimental groups may not yield unbiased causal estimates in the presence of nonrandom attrition. Comparing attrition rates between experimental groups can yield an unbiased estimate of the effect of the manipulation on attrition. However, if the manipulation is not independent of attrition (i.e., it is associated with differences in the rate or pattern of attrition), general comparisons between experimental groups on other outcomes are potentially subject to nonresponse bias.

For example, suppose that the constant presentation of the face induced high cognitive load relative to the strategic presentation, and this led to increased attrition in the constant face condition. Then the difference in attrition may be interpreted as a causal effect of the manipulation. However, differences in cognitive load would become increasingly biased with time as those who experience high cognitive load because of the constant presentation of the instructor's face leave the course. Although random assignment at the beginning of the experiment produces two comparable groups, differential attrition has the power to induce interdependence between the experimental manipulation and observable and unobservable covariates, yielding groups that are no longer comparable. We therefore report and interpret the findings with an appropriate degree of caution.

Results

Constant and strategic face comparisons for most dependent variables are provided in Table 2, with descriptive statistics, results of statistical tests, and effect size estimates. To investigate individual differences, learners who reported a preference for learning from visual information were compared with those with a preference for verbal information, which limited the sample to those who indicated a preference on the course survey ($N = 1{,}206$). Descriptive and inferential statistics for comparisons across learning preferences and video presentation styles are shown in Table 3.

Course grade and taking assessments. Course grades are a rich measure of a learners' recall and transfer learning, for they were based on a number of assessments throughout the course and an extensive final exam. Grades were not significantly different between video presentation styles on the basis of a nonparametric test, because grades were not normally distributed. This finding is consistent with prior work on the image principle and addresses Research Question 3.1 about differences in learning outcomes. Nevertheless, course grades were only available for learners who were motivated enough to complete assignments in the course.[3] We therefore compared learners' likelihood of taking assessments and found no significant difference between conditions. This sug-

[3] There was also no significant difference in learners' likelihood of earning a certificate in the course (awarded to learners who scored at least 70) across presentation styles. Course grades remained not significantly different when omitting the 89% of learners who scored a zero grade because they did not take assessments.

Table 2
Descriptive and Inferential Statistics for Assessment Taking, Course Grade, Cognitive Load, and Social Presence

Variable	Assessment taking	Course grade	Cognitive load	Social presence
Constant face	11.4% (720/5615)	8.9 (26.1)[a]	0.46 (0.19)[a]	0.57 (0.17)[a]
Strategic face	10.7% (660/5473)	8.6 (26.0)[a]	0.47 (0.20)[a]	0.59 (0.18)[a]
Test statistic	$\chi^2(1) = 1.10$[b]	$\chi^2(1) = 0.33$[c]	$z = 2.4$[d]	$z = 2.1$[d]
p Value	*n.s.*	*n.s.*	.016	.037
Effect size	1.06 odds ratio	$d = 0.01$	2.5% [0.5%, 4.6%][e]	2.2% [0.1%, 4.2%][e]
N	12,468	12,468	10,810 responses from 2,672 learners	5,965 responses from 2,169 learners

Note. *n.s.* = no significant effects were found. Effect sizes and statistical tests for cognitive load (CL) and social presence (SP) were derived from bootstrapped linear regression with a fixed effect for each video (CL) or week (SP) in the course.
[a] Mean (with standard deviation in parentheses). [b] Chi-square contingency table test. [c] Wilcoxon test. [d] Based on bootstrapped null distribution. [e] Bootstrapped relative change with 95% confidence interval.

gests that the strategic presence of the face influenced neither learning outcomes nor learners' motivation to engage with the course materials (Research Question 3.2).

We repeated the comparisons of course grades and assessment taking for learners who indicated a preference for either visual or verbal information to address research question Research Question 5c. There were no significant differences in assessment taking or course grades (see Table 3). All results remained statistically insignificant when restricting the sample to learners who had been more exposed to each treatment by watching more lecture videos.

Cognitive load. Although learners' self-reported cognitive load at the end of video lectures throughout the duration of the course, the following analysis was restricted to the first half of the course (Videos 1 to 9 in Figure 3) to reduce bias from differential attrition (cf. the **Statistical considerations** section). Learners in the strategic condition reported 2.5% ($z = 2.40$, $p = .016$, 95% confidence intervals[CI] = [0.5, 4.6]) higher cognitive load than learners in the constant condition (Research Question 1). We further examined individual differences in learning preferences and found an interaction effect between learners' preference and presentation style. The strategic face induced 20% higher cognitive load than the constant face for learners with a verbal preference. In contrast, learners with a visual preference reported 13% lower cognitive load with the strategic face than the constant face. These findings address Research Question 5a about individual differences in the effect of the strategic and constant presentation style on cognitive load and highlight the importance of learning preference as a moderating variable.

The average level of cognitive load varied considerably across lecture videos, which suggests that a substantial amount of variation in cognitive load was a result of properties of lecture videos other than the instructor's face. Not surprisingly, a closer examination of the cognitive load pattern revealed a strong correlation with video length, $r = .56$, $t(36) = 4.1$, $p < .001$. Average cognitive load for lecture videos in chronological order by experimental condition is illustrated in Figure 3.

Social presence. Learners reported the perceived social presence of the instructor at the end of a lecture video in each week of the course. Social presence increased by 9.3% between the first and last week of the course ($z = 30.00$, $p < .001$, 95% CI = [2.3, 16.4])—weekly means and standard errors are illustrated in Figure 4. As in the analysis of cognitive load, we restricted the following comparisons to the first half of the course (Weeks 1–5). Learners in the strategic face condition experienced 2.2% higher social presence than learners in the constant face condition (Research Question 2). Significantly higher social presence in Weeks 3, 4, and 7 for learners in the strategic than constant condition ($p < .05$) is shown in Figure 4. No significant differences in social presence were found at the beginning of the course, when ratings should be least biased by differential attrition. Learning preference was not a significant moderator of the effect of presentation style on social presence (Research Question 5b).

Attrition. Analysis of attrition in a learning environment where learners can join and leave at any time is a complex endeavor. Fortunately, the majority of learners joined in the first 3 weeks of the course (51% of learners started watching lectures in

Table 3
Descriptive and Inferential Statistics for Assessment Taking, Course Grade, Cognitive Load, and Social Presence by Learning Preference

Learning preference	Instructor face	Assessment taking	Course grade	Cognitive load	Social presence
Verbal	Constant	48% (100/208)	42.3 (45.0)[a]	0.43 (0.20)[a]	0.58 (0.18)[a]
Verbal	Strategic	51% (96/189)	45.8 (45.7)[a]	0.49 (0.19)[a]	0.59 (0.16)[a]
Visual	Constant	45% (178/393)	39.7 (44.8)[a]	0.46 (0.19)[a]	0.57 (0.18)[a]
Visual	Strategic	45% (186/416)	39.8 (45.2)[a]	0.43 (0.19)[a]	0.58 (0.18)[a]
Test statistic		$\chi^2(3) = 2.4$[b]	$\chi^2(3) = 2.2$[c]	$F(3, 487) = 3.0$[d]	$F(3, 812) = 0.6$[d]
p Value		*n.s.*	*n.s.*	.030	*n.s.*
N		1,206	1,206	4,509 responses from 810 learners	4,006 responses from 831 learners

[a] Mean (with standard deviation in parentheses). [b] Chi-square contingency table test. [c] Kruskal-Wallis test. [d] *F* statistic from hierarchical linear model to account for repeated measures with fixed effects for time units. *n.s.* = no significant effects were found.

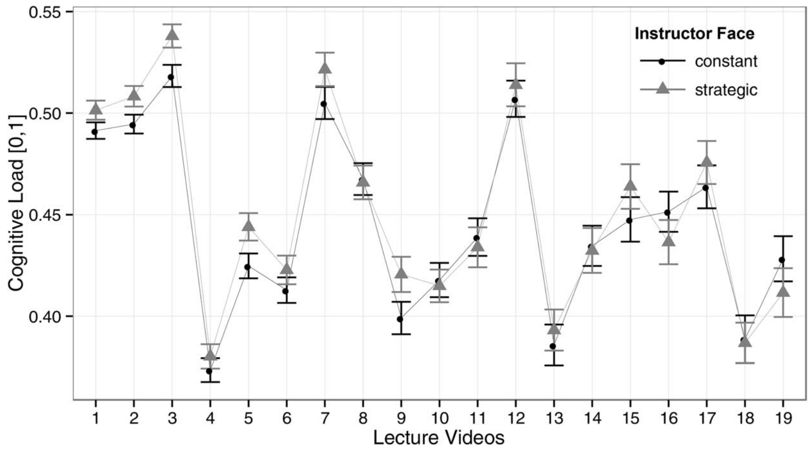


Figure 3. Cognitive load for each video in the course by presentation style showing higher cognitive load in the strategic condition early in the course and large variation in cognitive load, partly because of differences in video length (bootstrapped standard error bars).

the 1st week, 70% in the first 2 weeks, and 77% in the first 3 weeks of the course). There was no significant evidence for differential attrition between the constant and strategic condition for those who started watching in Week 1 ($z = 0.16$, $p = .85$), in Week 2 ($z = 1.60$, $p = .11$), or Week 3 ($z = 0.77$, $p = .44$). This addresses Research Question 4 about differential attrition.

Attrition in each condition is illustrated in Figure 5 by three pairs of Kaplan-Meier (KM) survival curves, each representing a subpopulation defined by video watching times. These KM curves show the proportion of learners who remained actively watching lectures at each point in the course. The three subpopulations represent learners who started watching in the 1st week, first 2 weeks, and first 3 weeks and watched at least one lecture after the 1st, 2nd, and 3rd week (*N*s = 2,903, 3,102, and 2,868, respectively).

A second survival analysis was conducted to investigate individual differences in attrition in response to the experimental manipulation (Research Question 5d). Learners who reported a preference for visual information were compared with those with a preference for verbal information, which limited the sample to those who completed the course survey and watched more than one lecture ($N = 1{,}250$). Because 96% of these learners watched a lecture during the first 3 weeks of the course, we restricted the analysis to learners who started watching lectures in the first 3 weeks and watched at least one lecture after the 3rd week. KM survival curves for each condition by learning preference are shown in Figure 6. The visible trends were tested for significance using Cox regression, which confirmed that attrition was 46% more likely in the strategic condition than in the constant condition for learners with a verbal preference, $z = 1.93$, $p = .05$; proportional hazard test: $\chi^2(1, N = 103) = 0.77$, $p = .38$. In contrast, for learners with a visual preference, the effect appeared to be reversed, albeit not statistically significant: Attrition in the constant condition was 22% higher than in the strategic condition, $z = 1.57$, $p = .12$. Learners with a verbal preference were overall 35% more likely to persist than those with a visual preference, $z = 2.59$, $p = .01$; proportional hazard test: $\chi^2(1, N = 357) = 0.11$, $p = .74$.

The largest difference in attrition was observed for learners with a visual preference in the strategic condition, who were 77% more

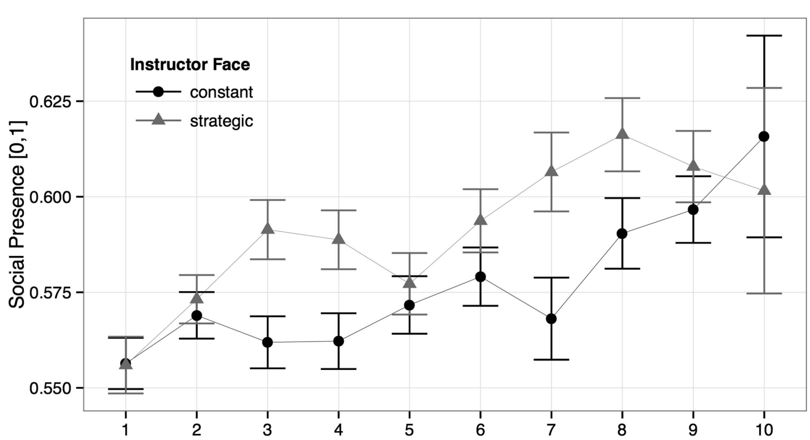


Figure 4. Social presence by presentation style, showing a steady increase over time in both conditions and higher social presence in the strategic condition in certain weeks (bootstrapped standard error bars).

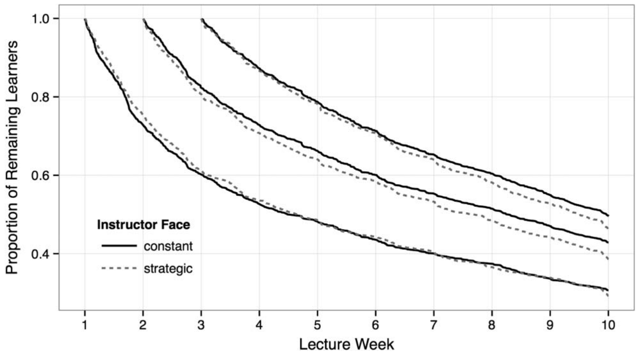


Figure 5. Kaplan-Meier survival curves showing similar rates of attrition for learners in each condition who started watching lectures in the ″first″, first 2, and first 3 weeks.

likely to drop out than learners with a verbal preference in the constant condition ($z = 3.49$, $p < .001$). This addresses Research Question 5b about individual differences in attrition.

Discussion

The effects of presenting the instructor's face strategically, instead of constantly, in video lectures were investigated in a longitudinal field experiment. The findings address a set of research questions that were developed on the basis of prior research in multimedia learning and related work. An overview of research questions and corresponding findings is provided in Table 4.

The strategic presentation—designed to better guide learners' attention and reduce the amount of distraction in exchange for lower levels of social presence—induced higher cognitive load and social presence relative to the constant presentation. This suggests that the primary feature of the strategic presentation style was not the presence and absence of the instructor's face but, rather, the appearance and disappearance of the instructor's face. This increased the instructor's face's salience, thereby inducing more social presence but also more distraction. Learning outcomes were not different across presentation styles, potentially because of the simultaneous increase in both social cues that promote learning and visual distractions that hinder learning.

Besides learning outcomes, we examined two measures of motivation that are most relevant in open learning environments—namely, learners' propensity to take assessments and persistence in the course. No main effect in assessment taking or attrition was found, but whether a learner preferred to learn from visual or verbal information moderated the effect of presentation styles on attrition. Learners with a verbal preference experienced lower levels of cognitive load and attrition in the constant than strategic condition. The effect was reversed for learners with a visual preference, albeit not significantly for attrition. Learning preference did not, however, moderate effects on learning outcomes, assessment taking, or social presence.

The results suggest that the design of the strategic presentation style was no improvement over the constant presentation, especially not for learners with a visual preference, who were substantially better served by the constant presentation style. The strategy underlying the strategic presentation seems to have failed through unintended consequences, which prompts future research on alternative implementations to balance social cues and visual distraction in video instruction. Learning preference, as an important individual difference, may simply reflect learners' ability to ignore the strategic face as a particularly distracting visual stimulus. This interpretation would account for the reversed effects on cognitive load for learners with different learning preferences.

Alternatively, the strategic presence of the instructor's face could induce extraneous cognitive load if learners fail to recognize and use the instructor's face as a visual signaling device. Learners with a verbal preference, for instance, may be less likely to recognize and use the presence of the instructor's face as a signaling device and may be more dependent on nonverbal

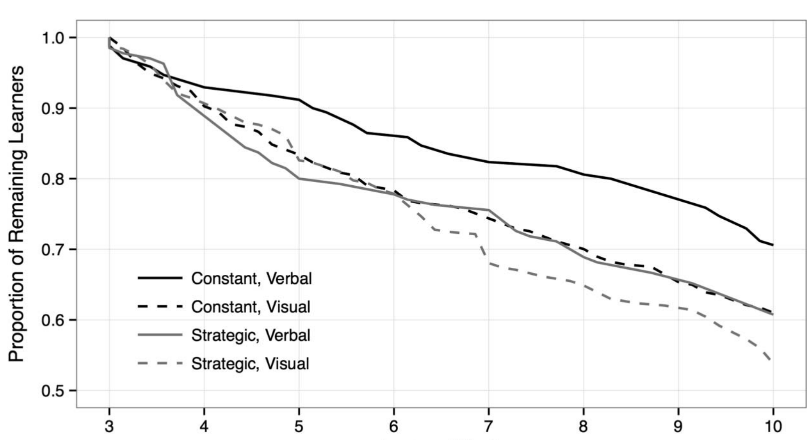


Figure 6. Kaplan-Meier survival curves showing differential attrition between presentation styles and learners' information preference. Learners with a verbal preference in the consistent face condition were 77% less likely to drop out than learners with a visual preference in the strategic condition.

Table 4
Overview of Research Questions (RQs) and Findings

RQ	Dependent variable	Main effect	Learning preference (RQ5)
RQ1	Cognitive Load	Strategic > Constant	Visual: Strategic > Constant
RQ2	Social Presence	Strategic > Constant	*n.s.*
RQ3.1	Learning Outcomes	*n.s.*	*n.s.*
RQ3.2	Assessment Taking	*n.s.*	*n.s.*
RQ4	Attrition	*n.s.*	Verbal: Strategic > Constant

Note. Individual differences by learning preference were examined for a subpopulation of learners who reported a preference. *n.s.* = no significant effects were found.

cues to supplement verbal understanding, leading to higher levels of cognitive load and attrition in the strategic condition.

General Discussion

With an observational study and an experiment of unprecedented magnitude in the multimedia learning literature, we aimed to extend prior research on the image and personalization principle (Mayer, 2001) by examining learners' behavior, attitudes, and learning outcomes in two large online courses.

Summary of Findings

Study 1 highlighted the fact that although the majority of learners preferred to watch video lectures with the instructor's talking head, a substantial number of learners opted for lectures without it or switched between versions. Learners' rationales for choosing a presentation style, on the one hand, highlighted the benefits of social and other nonverbal cues from the instructor, which reportedly helped learners' focus and feel more connected. On the other hand, it highlighted the downsides of incurring extraneous cognitive load by processing the mostly irrelevant talking head. This inspired the development of the strategic presentation style as an informed compromise between the two extremes, which was compared with constantly presenting the face in a field experiment (Study 2). The results were consistent with prior research on multimedia learning, with no differences in learning outcomes across presentation styles. Overall, the strategic presentation style was inferior, especially for learners with a visual learning preference.

Learning preference is highlighted as an individual difference to consider in the design of multimedia instruction (in the worst case scenario, one group of learners was 77% more likely to drop out of the course). Homer et al. (2008) previously identified learners' visual cognitive style as a moderator of the effect of showing the instructor's face on cognitive load.

Implications

The present findings have implications for theory and practice. The strategic presentation style was inspired by our findings from Study 1 and derived from theory and prior empirical evidence, which has highlighted the benefits of signaling, coherence, and personalization (Clark & Mayer, 2011). The presence and absence of the instructor's face in the strategic condition was intended to guide learners' attention (signaling) and reduce the cognitive burden of split visual attention (coherence) while preserving occasional social and other nonverbal cues from the instructor (personalization). However, the salience of the instructor's face seemed to be increased in the strategic condition, even though the instructor's face was only visible between one- and two-thirds of the time that it was visible in the constant condition. Instructional designers should be aware of this somewhat counterintuitive effect of strategically presenting the instructor.

Another insight for the practice of producing and learning with video lectures is that although the majority of learners prefer seeing the instructor's face and believe that it helps them concentrate and understand the instructor, a large percentage of learners instead prefer not to see the instructor because they find it distracting. It is unclear whether learners' beliefs translate into better learning outcomes and to what extent learners can make an optimal choice in this respect. We, therefore, suggest follow-up research to address these open questions, including the efficacy of adaptive or learner-controlled interfaces in multimedia learning.

The present findings offer two theoretical insights, one regarding the effect of social cues on persistence in open learning environments, the other pertaining to learning preference as an individual difference in multimedia learning. The personalization principle is aimed at motivating learners to exert effort to learn by adding social cues to instruction. Going beyond effects on learning outcomes, our longitudinal field experiment used two direct behavioral measures of the effect of increasing the salience of social cues in an open learning environment where learners could choose whether to take assessments and continue watching lectures in the course. Learners in the strategic condition reported higher levels of social presence than those in the constant condition but were, in fact, not any more likely to take assessments or persist in the course (no significant differences in social presence were found in the moderation analysis for learning preference). This apparent inconsistency could stem from the small manipulation of social presence and the concurrent manipulation of cognitive load. It highlights the potential for future research on the effects of personalization in open learning environments, specifically on behaviors that are necessary for learning to occur (e.g., persisting in a course, engaging with learning materials).

The present study provides strong evidence for the importance of learning preference as an individual difference for effects on cognitive load and attrition. In a large body of literature investigating the role of individual differences in multimedia learning (e.g., Kalyuga, 2005; Kalyuga et al., 2003; Kirby, 1993; Kirby, Moore, & Schofield, 1988; Mayer & Sims, 1994; Plass, Chun, Mayer, & Leutner, 1998; Rayner & Riding, 1997), learning preference has been examined repeatedly but is not as established as spatial ability (Mayer & Sims, 1994) or prior knowledge (Kalyuga, 2005). The present findings highlight learning preference as a critical individual difference to consider in the design of multimedia instruction. Further investigation of the role of information preference will require the development of an improved measure.

Limitations and Future Work

Despite the magnitude of this study, the generalizability of the present findings is limited in a number of ways. In particular, we acknowledge the weakness of our study design, resulting from the specificity of the task, duration, and population and the lack of control in the field study. MOOCs were a good choice for studying the behavior of many learners in a specific environment for a mostly highly educated adult learner population. Courses run for a specific period, and researchers have no control over learners' behavior, which has posed methodological challenges for data analysis, especially for the analysis of attrition. The study design bears challenges for obtaining unbiased estimates of causal quantities due to attrition (discussed in the Method section of Study 2) and self-selection to complete optional self-report measures.

Nevertheless, there are certain advantages to conducting field research, because it allows for natural observation of behavior in real-life learning scenarios, which in the case of this study extended for up to 10 weeks. In contrast to laboratory experiments, field studies provide higher levels of external validity in exchange for lower levels of control over the environment and procedure. Moreover, detecting smaller effects and conducting exploratory subgroup analyses requires large sample sizes. Future work could further investigate the findings from our studies in an environment with more experimental control to triangulate the reliability of our results and identify potential border conditions.

An important limitation of this study is the homogeneity of lecture videos in which the presentation style of the instructor's face was manipulated. All lectures showed lecture slides with text and occasional graphics, and the instructor was the same in all videos. Although many published lecture videos follow a similar format, many lecture videos on topics in computer science and mathematics, for instance, feature extensive real-time writing and annotation by the instructor. Different styles of video lectures might benefit from different dynamics for the presentation of the instructor's face. Therefore, the present findings should be replicated before applying insights from this study to lecture videos that use very different instructional methods.

Although Study 1 focused on the choice between videos with or without the instructor's face, Study 2 compared the strategic presentation only with the constant presentation, omitting the condition without the instructor's face. A no-face condition would have been a natural and sound addition to Study 2, but we chose to omit this condition because we were not comfortable assigning thousands of learners to an unconventional presentation style that was also found to be less preferred in Study 1 and a pretest. In this pretest, several online learners who were assigned to the no-face condition complained about difficulties concentrating during the lecture video. An experiment comparing all three conditions would be an appropriate follow-up to replicate as well as extend the current study. The strategic face version could be designed using a more rigorous method by using multiple coders (instead of a single coder) to identify segments in which the face should be removed with high intercoder reliability.

This study relied, in part, on self-reported measures that do not necessarily reflect corresponding internal states of learners, even in the absence of measurement error. Self-reported cognitive load, for instance, represents learners' own judgments of the amount of effort they invested relative to their expectations and past experiences. In the context of multimedia instruction, subjective measures can be very effective in quantifying facets of a learner's experience with lecture materials. Learning preference, which was measured using a single item, may have failed to gain traction in the field because of challenges in its measurement. The most common scale used for distinguishing visual from verbal cognitive style is A. Richardson's (1977) verbalizer–visualizer scale, which has been criticized for its unfavorable psychometric properties and construct validity (Kozhevnikov et al., 2002, 2005). The present study neither attempted to offer a revised measure of learning preference nor used Richardson's visualizer–verbalizer scale.

Future research could disentangle the potential benefits and costs of showing the instructor's face in multimedia learning in a more controlled setting. What are the relative contributions of providing social cues and supporting comprehension with a nonverbal communication channel to enable lip reading? This question could be investigated by comparing the video of the instructor with a static image, which provides a social cue but no nonverbal communication.

Another direction for future work could be to examine the effect of allowing learners to choose a presentation style compared with assigning them one, looking at whether learners tend to choose optimally in terms their subsequent learning outcomes. Because of the absence of a no-choice control condition and random assignment, this research question could not be addressed in Study 1. On the basis of the present findings on the moderating effect of learning preference, it would be worth investigating whether learners would choose the presentation style that is expected to induce lower cognitive load and attrition for them.

Conclusion

The prevalence of multimedia learning increases as traditional instruction becomes more and more augmented by technology. A large body of literature, with contributions dating back to the 1970s, has established guidelines based on empirical studies and learning theories on how to design instruction to promote cognitive processes that enable learning to occur. As technology keeps advancing, new opportunities arise for the design of multimedia instruction, with a handful that could, in fact, be leveraged to enhance learning. The image principle (Mayer, 2001) exemplifies a new technology (picture-in-picture video) that appears to be a helpful feature and is endorsed by learners but does not actually enhance learning.

The current work provides a novel perspective on the image principle by observing effects in two realistic, longitudinal field studies. The present findings suggest that although social cues may not enhance learning per se, they may affect learners' motivation to persist in a course, which is expected to benefit learners in the long run. Moreover, a one-size-fits-all strategy for the presentation of the instructor's face was found to adversely affect some learners by inducing higher cognitive load and attrition, depending on their learning preference. In light of the rapid growth of multimedia instruction, the current findings call for further research on the potential benefits of adaptive or learner-controlled systems to account for individual differences between learners.

References

Angrist, J. D., & Pischke, J. S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University Press.

Atkinson, R. K., Mayer, R. E., & Merrill, M. M. (2005). Fostering social agency in multimedia learning: Examining the impact of an animated agent's voice. *Contemporary Educational Psychology, 30,* 117–139. http://dx.doi.org/10.1016/j.cedpsych.2004.07.001

Baddeley, A. D. (1986). *Working memory*. Oxford, England: Oxford University Press.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4,* 829–839. http://dx.doi.org/10.1038/nrn1201

Bransford, J., Brown, A., & Cocking, R. (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.

Brünken, R., Plass, J. L., & Leutner, D. (2003). Direct measurement of cognitive load in multimedia learning. *Educational Psychologist, 38,* 53–61. http://dx.doi.org/10.1207/S15326985EP3801_7

Clark, R. C., & Mayer, R. E. (2011). *E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning*. San Francisco: Pfeiffer. http://dx.doi.org/10.1002/9781118255971

Cui, G., Lockee, B., & Meng, C. (2013). Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and Information Technologies, 18,* 661–685.

Dawson, M. E., Schell, A. M., Beers, J. R., & Kelly, A. (1982). Allocation of cognitive processing capacity during human autonomic classical conditioning. *Journal of Experimental Psychology: General, 111,* 273–295. http://dx.doi.org/10.1037/0096-3445.111.3.273

Dede, C. (2009, January 2). Immersive interfaces for engagement and learning. *Science, 323,* 66–69. http://dx.doi.org/10.1126/science.1167311

Dekle, D. J., Fowler, C. A., & Funnell, M. G. (1992). Audiovisual integration in perception of real words. *Perception & Psychophysics, 51,* 355–362. http://dx.doi.org/10.3758/BF03211629

Felder, R. M., & Spurlin, J. (2005). Applications, reliability and validity of the index of learning styles. *International Journal of Engineering Education, 21,* 103–112.

Ginns, P. (2006). Integrating information: A meta-analysis of the spatial contiguity and temporal contiguity effects. *Learning and Instruction, 16,* 511–525. http://dx.doi.org/10.1016/j.learninstruc.2006.10.001

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. (2001). Explaining math: Gesturing lightens the load. *Psychological Science, 12,* 516–522. http://dx.doi.org/10.1111/1467-9280.00395

Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer–mediated conferencing environment. *American Journal of Distance Education, 11,* 8–26.

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM Conference on Learning at Scale* (pp. 41–50). New York: Association for Computing Machinery. http://dx.doi.org/10.1145/2556325.2566239

Hapeshi, K., & Jones, D. (1992). Interactive multimedia for instruction: A cognitive analysis of the role of audition and vision. *International Journal of Human–Computer Interaction, 4,* 79–99. http://dx.doi.org/10.1080/10447319209526029

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology, 52,* 139–183.

Homer, B. D., Plass, J. L., & Blake, L. (2008). The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior, 24,* 786–797. http://dx.doi.org/10.1016/j.chb.2007.02.009

Horton, D., & Wohl, R. R. (1956). Mass communication and para-social interaction; observations on intimacy at a distance. *Psychiatry, 19,* 215–229.

Kalyuga, S. (2005). *Prior knowledge principle in multimedia learning*. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 325–337). New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816819.022

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38,* 23–31. http://dx.doi.org/10.1207/S15326985EP3801_4

Kartal, G. (2010). Does language matter in multimedia learning? Personalization principle revised. *Journal of Educational Psychology, 102,* 615–624. http://dx.doi.org/10.1037/a0019345

Kirby, J. R. (1993). Collaborative and competitive effects of verbal and spatial processes. *Learning and Instruction, 3,* 201–214. http://dx.doi.org/10.1016/0959-4752(93)90004-J

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary Educational Psychology, 13,* 169–184. http://dx.doi.org/10.1016/0361-476X(88)90017-3

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect. In *Proceedings of the annual SIGCHI Conference on Human Factors in Computing Systems* (pp. 2095–2102), New York: Association for Computing Machinery. http://dx.doi.org/10.1145/2556288.2557207

Kizilcec, R. F., Piech, C., & Schneider, E. (2013). Deconstructing disengagement: Analyzing learner subpopulations in massive open online courses. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 170–179). New York: Association for Computing Machinery. http://dx.doi.org/10.1145/2460296.2460330

Koller, D., & Ng, A. (2014, March 24). Welcome Rick Levin as CEO of Coursera. Message posted to http://blog.coursera.org/post/80601201906/welcome-rick-levin-as-ceo-of-coursera

Kozhevnikov, M., Hegarty, M., & Mayer, R. E. (2002). Revising the visualizer–verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction, 20,* 47–77. http://dx.doi.org/10.1207/S1532690XCI2001_3

Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: A new characterization of visual cognitive style. *Memory & Cognition, 33,* 710–726. http://dx.doi.org/10.3758/BF03195337

Lyons, A., Reysen, S., & Pierce, L. (2012). Video lecture format, student technological efficacy, and social presence in online courses. *Computers in Human Behavior, 28,* 181–186.

MacLeod, A., & Summerfield, Q. (1987). Quantifying the contribution of vision to speech perception in noise. *British Journal of Audiology, 21,* 131–141. http://dx.doi.org/10.3109/03005368709077786

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Educational Psychology, 81,* 240–246.

Mayer, R. E. (2001). *Multimedia Learning*. Cambridge University Press.

Mayer, R. E. (2005). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 183–200). New York: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511816819.013

Mayer, R. E., Dow, G. T., & Mayer, S. (2003). Multimedia learning in an interactive self-explaining environment: What works in the design of agent-based microworlds? *Journal of Educational Psychology, 95,* 806–812. http://dx.doi.org/10.1037/0022-0663.95.4.806

Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93,* 187–198. http://dx.doi.org/10.1037/0022-0663.93.1.187

Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology, 95,* 833–846. http://dx.doi.org/10.1037/0022-0663.95.4.833

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38,* 43–52. http://dx.doi.org/10.1207/S15326985EP3801_6

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86,* 389–401. http://dx.doi.org/10.1037/0022-0663.86.3.389

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology, 95,* 419–425. http://dx.doi.org/10.1037/0022-0663.95.2.419

McGurk, H., & MacDonald, J. (1976, December 23). Hearing lips and seeing voices. *Nature, 264,* 746–748. http://dx.doi.org/10.1038/264746a0

McLaren, B. M., DeLeeuw, K. E., & Mayer, R. E. (2011). A politeness effect in learning with Web-based intelligent tutors. *International Journal of Human–Computer Studies, 69,* 70–79. http://dx.doi.org/10.1016/j.ijhcs.2010.09.001

Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97. http://dx.doi.org/10.1037/h0043158

Moreno, R., & Mayer, R. E. (2000). Engaging students in active learning: The case for personalized multimedia messages. *Journal of Educational Psychology, 92,* 724–733. http://dx.doi.org/10.1037/0022-0663.92.4.724

Moreno, R., Mayer, R. E., Spires, H. A., & Lester, J. C. (2001). The case for social agency in computer-based teaching: Do students learn more deeply when they interact with animated pedagogical agents? *Cognition and Instruction, 19,* 177–213. http://dx.doi.org/10.1207/S1532690XCI1902_02

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56,* 81–103. http://dx.doi.org/10.1111/0022-4537.00153

Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84,* 429–434. http://dx.doi.org/10.1037/0022-0663.84.4.429

Paas, F. G. W. C., & van MerriËnboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology, 86,* 122–133. http://dx.doi.org/10.1037/0022-0663.86.1.122

Packham, G., Jones, G., Miller, C., & Thomas, B. (2004). E-learning and retention: Key factors influencing student withdrawal. *Education & Training, 46,* 335–342. http://dx.doi.org/10.1108/00400910410555240

Papert, S. (1993). *The children's machine: Rethinking school in the age of the computer*. New York: Basic Books.

Plass, J. L., Chun, D. M., Mayer, R. E., & Leutner, D. (1998). Supporting visual and verbal learning preferences in a second-language multimedia learning environment. *Journal of Educational Psychology, 90,* 25–36. http://dx.doi.org/10.1037/0022-0663.90.1.25

Rayner, S., & Riding, R. (1997). Towards a categorisation of cognitive styles and learning styles. *Educational Psychology, 17,* 5–27. http://dx.doi.org/10.1080/0144341970170101

Reeves, B., & Nass, C. I. (1996). *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge University.

Richardson, A. (1977). Verbalizer–visualizer: A cognitive style dimension. *Journal of Mental Imagery, 1,* 109–125.

Richardson, J. C., & Swan, K. (2003). Examining social presence in online courses in relation to students' perceived learning and satisfaction. *Journal of Asynchronous Learning Networks, 7,* 68–88.

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *The Internet and Higher Education, 6,* 1–16. http://dx.doi.org/10.1016/S1096-7516(02)00158-6

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. New York: Wiley.

Strobach, T., Liepelt, R., Schubert, T., & Kiesel, A. (2012). Task switching: Effects of practice on switch and mixing costs. *Psychological Research, 76,* 74–83. http://dx.doi.org/10.1007/s00426-011-0323-x

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12,* 257–285. http://dx.doi.org/10.1207/s15516709cog1202_4

Sweller, J., van MerriËnboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10,* 251–296. http://dx.doi.org/10.1023/A:1022193728205

Tranel, D., Fowles, D. C., & Damasio, A. R. (1985). Electrodermal discrimination of familiar and unfamiliar faces: A methodology. *Psychophysiology, 22,* 403–408. http://dx.doi.org/10.1111/j.1469-8986.1985.tb01623.x

Tu, C.-H., & McIsaac, M. (2002). The relationship of social presence and interaction in online classes. *American Journal of Distance Education, 16,* 131–150. http://dx.doi.org/10.1207/S15389286AJDE1603_2

van MerriËnboer, J. J. G., Kester, L., & Paas, F. (2006). Teaching complex rather than simple tasks: Balancing intrinsic and germane load to enhance transfer of learning. *Applied Cognitive Psychology, 20,* 343–352. http://dx.doi.org/10.1002/acp.1250

Waldrop, M. M. (2013, March 14). Online learning: Campus 2.0. *Nature, 495,* 160–163. http://dx.doi.org/10.1038/495160a

Wang, N., Johnson, W. L., Mayer, R. E., Rizzo, P., Shaw, E., & Collins, H. (2008). The politeness effect: Pedagogical agents and learning outcomes. *International Journal of Human–Computer Studies, 66,* 98–112. http://dx.doi.org/10.1016/j.ijhcs.2007.09.003

Willging, P. A., & Johnson, S. D. (2004). Factors that influence students' decision to drop out of online courses. *Journal of Asynchronous Learning Networks, 8,* 105–118.

Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1–10). New York: Association for Computing Machinery A. http://dx.doi.org/10.1145/1240624.1240626

Received May 3, 2014
Revision received October 10, 2014
Accepted October 11, 2014 ■