Accepted Manuscript

Instructor presence effect: Liking does not always lead to learning

Kristin E. Wilson, Mark Martinez, Caitlin Mills, Sidney D'Mello, Daniel Smilek, Evan F. Risko

PII: S0360-1315(18)30064-2

DOI: 10.1016/j.compedu.2018.03.011

Reference: CAE 3321

To appear in: *Computers & Education*

Received Date: 17 June 2017

Revised Date: 9 March 2018

Accepted Date: 10 March 2018

Please cite this article as: Wilson K.E., Martinez M., Mills C., D'Mello S., Smilek D. & Risko E.F., Instructor presence effect: Liking does not always lead to learning, *Computers & Education* (2018), doi: 10.1016/j.compedu.2018.03.011.

Instructor presence effect: Liking does not always lead to learning

Kristin E. Wilson*[1], Mark Martinez[1], Caitlin Mills[2], Sidney D'Mello[2], Daniel Smilek[1] & Evan F. Risko[1]

University of Waterloo, Department of Pyschology[1]
University of Notre Dame, Department of Psychology[2]

Submitted June 17, 2017

*Corresponding Author:
Dr. Kristin E. Wilson
Department of Psychology
University of Waterloo
200 University Ave
Waterloo, ON, Canada
N2L3G1
kristin.wilson@uwaterloo.ca

Instructor presence effect: Liking does not always lead to learning

(Manuscript Word Count: 10,436; Abstract Word Count: 243)

Submitted: June 17, 2017

Abstract

Online education provides the opportunity to present lecture material to students in different formats or modalities, however there is debate about which lecture formats are best. Here, we conducted four experiments with 19 to 68 year old online participants to address the question of whether visuals of the instructor in online video lectures benefit learning. In Experiments 1 (N = 168) and 2 (N=206) participants were presented with a lecture in one of three modalities (audio, audio with text, or audio with visuals of the instructor). Participants reported on their attentiveness – mind wandering (MW) – throughout the lecture and then completed a comprehension test. We found no evidence of an advantage for video lectures with visuals of the instructor in terms of a reduction in MW or increase in comprehension,. In fact, we found evidence of a comprehension cost, suggesting that visuals of instructors in video lectures may act as a distractor. In Experiments 3 (N=88) and 4 (N=109) we explored learners' subjective evaluations of lecture formats across 4 different lecture formats (audio, text, audio+text, audio+instructor, audio+text+instructor). The results revealed learners not only find online lectures with visuals of the instructor more enjoyable and interesting, they believe this format most facilitates their learning. Taken together, these results suggest visuals of the instructor potentially impairs comprehension, but learners prefer and believe they learn most effectively with this format. We refer to as the *Instructor Presence Effect* and discuss implications for multimedia learning and instructional design.

Keywords: Online lectures; mind wandering; learner preferences; judgment of learning; seductive details

1. Introduction

The growth of online education in the digital age has been rapid and profuse, with many universities and colleges now offering blended and fully online courses in growing numbers (American Association of Community Colleges, 2015; Martel, 2015). While there are advantages to liberating education from the traditional bricks and mortar classroom, there are some unique challenges. For example, online lecture drop-rates are high, with about 50% of learners having dropped lectures (i.e., turned the video off) within the first 5 minutes (Kim et al., 2014). Although competing with other demands (e.g., social media) is a significant challenge, simply attending to a video lecture is also difficult. Even when learners make it through an entire video lecture, it is unclear how much of that lecture was actively attended, as inattention or mind wandering (MW, i.e., engaging in task-unrelated, internally generated thoughts) increases as a function of time (Farley, Risko, & Kingstone, 2013; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Risko, Buchanan, Medimorec, & Kingstone, 2013). What can be done to make online lectures more engaging? A reasonable place to start is by considering the different ways in which the lecture material is presented in online courses. For instance, some lectures are presented as audio only recordings (with or without text), while others include visuals of the instructor (with or without text). In this paper we examine whether adding a video of an instructor to online lectures might improve attentional engagement, learning, and subjective evaluations (e.g., preferences, judgments of learning).

1.1 Possible benefits and costs of adding a video of an instructor to online lectures

There are several reasons to believe that adding video of an instructor to online lectures might *facilitate* attentional engagement with the lecture material, thus enhancing learning and enjoyment. One reason concerns the possibility that adding an instructor might draw attention away from mind wandering and attract attention to the lecture material. As noted above, there is a growing body of evidence showing that mind wandering is common during video lecture viewing and that it is associated with on worse learning outcomes (Farley et al., 2013; Risko et al., 2012, 2013). This is primarily because attentional resources required for processing lecture material are diverted during mind wandering. It is also well known that attention can be captured by salient stimuli and events in the environment, such as socially relevant information (Langton & Bruce, 1999; Langton, Watt, & Bruce, 2000) facial features (Langton, Law, Burton, & Schweinberger, 2008; Jan Theeuwes & Van der Stigchel, 2006) and motion (Abrams & Christ, 2003). Accordingly, by virtue of their social significance, instructors' facial features and motion while teaching might act as salient cues that trigger an automatic exogenous shift (Hickey, McDonald, & Theeuwes, 2006; Theeuwes & Burger, 1998; Jan Theeuwes, 2004) in the allocation of resources from mind wandering to the lecture material. The improvement in attentional focus created by the addition of an

instructor might ultimately improve learning and may contribute to a heightened subjective sense of engagement with the lecture. The subjective sense of engagement or feeling of attention being drawn towards the instructor may be interpreted by students as an internal index of learning, thus potentially influencing their judgments of learning.

Another reason that adding a video of an instructor to online lectures might facilitate attentional engagement concerns the relation between attentional engagement and the cognitive load imposed by a task. When attention is not sufficiently engaged by an external task people tend to mind wander, which is operationalized as a state of being 'off-task', directing attention away from external stimuli/content and towards internally generated thoughts (Smallwood, 2013; Smallwood & Schooler, 2006). It has been shown that mind wandering increases as the resource requirements of a task decrease (i.e., the cognitive load) (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; McVay, Kane, & Kwapil, 2009; Smallwood & Schooler, 2006). Resources not used by the task often tend to be allocated to internal thought. As an example, Kopp and D'Mello (2016) compared the proportion of time spent MW (estimated from the proportion of positive responses to thought probes throughout the task) in online learners when learning materials (literary passages) were of higher cognitive demand (i.e., dual-processing channels, visual and auditory) relative to lower cognitive demand (i.e., single-channel, audio only or text only). The results revealed that MW was lowest in the dual-mode condition compared to audio-only, computer-paced reading, and self-paced reading conditions for fast readers. This suggests that presenting learners with multimodal material that increases intrinsic cognitive load may facilitate learning through reducing internally generated distraction or MW. It is possible that a visual of the instructor during a video lecture may result in additional cognitive/attentional processing (e.g., attending to and interpreting instructor gestures), increasing intrinsic cognitive load and reducing MW.

Conversely, adding an instructor might *hinder* attentional engagement with lecture content due to attention-capture effects of adding an instructor to the lecture. Although the instructor might be a salient cue that draws attentional resources away from mind wandering, these resources might be devoted to visual processing of the instructor rather than the relevant lecture material. Distinguishing between attentional engagement with relevant and irrelevant content may be difficult and often confused by learners. As such, focus on the instructor may be interpreted by learners as engagement with lecture material when in fact it reflects engagement with salient visual features that are irrelevant to learning. On this possibility, adding an instructor might reduce mind wandering, but increase attention to irrelevant instructor characteristics. This should have little impact on learning (resources from mind wandering are just shifted to irrelevant information) while leading to subjective evaluations of greater engagement and

learning. In this case, we might expect to see a dissociation among mind wandering, learning, and subjective evaluations.

Let us also reconsider the possibility that increasing cognitive load might have broad sweeping benefits for attentional engagement, learning, and subjective experience (e.g., enjoyment and self-evaluations or beliefs about learning). Although there is evidence that increasing cognitive load may decrease mind wandering, there is also research suggesting that increasing cognitive load negatively impacts learning (Chaiken & Eagly, 1976; Mayer, Heiser, & Lonn, 2001; Mousavi, Low, & Sweller, 1995; Otondo, Van Scotter, Allen, & Palvia, 2008). Specifically, the added cognitive load of a visual of an instructor might reduce mind wandering, but because the cognitive load is extraneous (i.e., not directly related to the learning material), it might also hinder learning, with unknown effects on subjective experience (possibly again increasing the subjective sense of engagement with the material). Indeed, it is not uncommon for learners to report a preference for learning material that contains irrelevant 'seductive' details, but unbeknownst to the learner, processing this extraneous information wastes capacity-limited resources and impairs learning (Harp & Mayer, 1998; Mayer et al., 2001; Mousavi et al., 1995; Paas, Renkl, & Sweller, 2003; Sung & Mayer, 2012).

Based on the foregoing theoretical considerations, adding a video of an instructor to online lectures could have a variety of consequences on attentional engagement, learning and subjective evaluations. It also becomes clear that a resolution of the various possibilities would require a comprehensive assessment of attentional engagement (indexed by levels of mind wandering), learning (indexed by assessments of lecture material), and subjective evaluations (judgment of learning, preference etc.). At present, however, such a comprehensive assessment is lacking. That being said, a few studies have begun to chip away at these issues, which we turn to next.

1.2 The available evidence

There is some evidence that using tools that enhance social agency or social presence in online lectures fosters learning (Atkinson, Mayer, & Merrill, 2005; Mayer, Sobko, & Mautone, 2003; Moreno, Mayer, Spires, & Lester, 2001). Including visuals of the instructor in online courses is one way of increasing social agency and presence in online lectures and there is some evidence that this enhances learners' subjective sense of engagement. For instance, using a correlational design, Gunawardena and Zittle (1997) found that students' ratings of instructor social presence correlated with students' self-reports of course satisfaction. Another study found that students' level of self-reported engagement in an online course increased as instructor-personalized material (i.e., voice or visuals/video of instructor) was cumulatively increased over successive lectures, from power point alone, to power point with audio, to power point with

video of the instructor (Mandernach, 2009). Although students reported a greater degree of engagement with increases in instructor-personalized material, it had no measurable impact on student grades. A critical problem with this study, however, was that the video of the instructor was added later in the course for all students and so increases in lecture-personalization were confounded with the number of lectures that had already been presented. As such, it is possible that the instructor-personalized material did have a positive impact on grades, which was essentially cancelled out by an underlying increase in difficulty of the material over time.

Using a more controlled manipulation of instructor presence, Kizilcec, Bailenson, and Gomez (2015) explored how the presence of an instructor impacts learners' preferences for, and choice of, various presentation formats in a Massive Open Online Course. Learners were presented with two different lecture formats in the first week of classes: One format did not contain visuals of the instructor and the other did. In the second week learners had the option to choose which lecture they would watch and at the end of the course learners were asked to report on their preference for the two different lecture formats. The results confirm previous findings, as the majority of participants *reported* a preference for the video with the instructor with only 16% reporting a preference for the video without the visual of the instructor. When the researchers investigated which video students chose to watch, however, 35% of learners consistently *chose to watch* the video that did not contain the instructor, with another 8% choosing to watch both. We should be cautious in our interpretation of these results, however, as there was a significant order effect such that participants tended to watch the video type that was presented to them first. Furthermore, only learners who watched every lecture in the first week were included in the analysis, which resulted in the exclusion of nearly 90% of learners enrolled in the course, leading to a possible self-selection bias.

An eye-tracking study reported by Kizilcec, Papadopoulos, and Sritanyaratana (2014) revealed that when participants are presented with visuals of the instructor in a video lecture, visual attention (measured via eye fixations) is drawn to the instructor. In this study, learners spent almost half the time looking at the instructor when s/he was included in the video, yet the presence versus absence of the instructor, had no impact on retention or memory for lecture material. In light of previous research, showing that participants substantially change their behavior when they know their eye movements are being monitored (Risko & Kingstone, 2011), it is possible that the use of the eye-tracker in Kizilcec et al.'s (2014) study had an influence on attention. It is also unclear how attentional engagement may vary across laboratory and more authentic online learning contexts, since distractions are different across these two contexts. These considerations pose some possible limitations on the generalizability of these findings.

There is some evidence that presenting learners with video of an instructor may interfere with learning, which may be the result of increased cognitive load and visual distraction. Fiorella and Mayer

(2016) tested the impact of video lectures, with and without visuals of instructors' faces/bodies on learning. Specifically they aimed to address the question of whether learners' understanding of a concept (the Doppler effect) could be facilitated by seeing the instructor explain and draw a diagram of the effect. The results provided evidence that watching an instructor *draw a diagram* benefits learning outcomes (transfer of concepts), relative to when the instructor merely stands beside the complete diagram and explains it (with and without pointing/gestures), particularly for those with low prior knowledge. This result suggests that showing the instructor drawing a diagram can facilitate learning, which might operate through a mechanism of cuing attention to relevant information (i.e., the diagram). In a final experiment one group of participants was presented with a video containing a visual of the instructor (head and torso) as she drew and explained the diagram while another group watched a video showing only the instructor's hand as she drew and explained the diagram. This final experiment revealed that the visual of the instructor was associated with *a learning cost* relative to when just the hand was visible, irrespective of prior knowledge. This learning cost suggests that visuals of the instructor may engage and capture visual attention, guiding attention towards irrelevant visual characteristics of the instructor, potentially distracting learners from conceptual content in the lecture.

The notion that visuals of an instructor may act as a salient and seductive distractor is supported by the eye tracking data mentioned earlier, suggesting that learners attention is attracted to visuals of the instructor (Kizilcec et al., 2014). Under normal viewing conditions (without an eye-tracker), in a more authentic online learning environment, attentional resources that otherwise would be dedicated to processing conceptual content in an online lecture may be allocated to processing visual features of the instructor and /or tracking the instructor's movements, ultimately impairing learning. There is some anecdotal evidence of this in Kizilcec et al's., (2015) study, where some participants' reports (free-form written feedback on their experience of the different lectures) allude to the instructor being distracting. In a follow-up experiment, Kizilcec et al., (2015) attempted to overcome this issue of possible distraction by creating a lecture that had the instructor appear 'strategically' and periodically throughout the lecture. This viewing condition, relative to a lecture with a constant visual of the instructor, however, was associated with subjective reports of *higher* cognitive load, suggesting greater distraction. Although this study did not contain an objective measure of attention, the higher reports of cognitive load with intermittent presentation of the instructor is consistent with the notion that intermittent visuals of the instructor may act as even more of a distractor. Prior research shows that inhibitory control over processing information deemed irrelevant is strengthened over time (Dixon, Ruppel, Pratt, & De Rosa, 2009). Thus, in the constant instructor condition, it may become easier to ignore the instructor over time, reducing cognitive

load. The periodically appearing instructor, on the other hand, likely makes it difficult to habituate to this visual stimulus, leading to stronger attentional capture.

1.3 Present Investigation

Taken together, there is some evidence that visuals of the instructor may attract the gaze of learners and capture attention in lectures presented in the laboratory. However, it is unclear what impact this has on learners' attention, comprehension, and preferences in online learning environments relative to other modalities of presenting lecture material. A rigorous approach to this question requires the controlled manipulation of lecture format in order to establish a cost/benefit relative to alternative lecture formats and comprehensive measurement of not only learning (e.g., comprehension) and learners' subjective evaluations, but also a measure of the control of attention. Given that video recorded lectures with visuals of the instructor are multi-channel (audio/visual), it is important that this lecture format be contrasted with another multi-channel format that does not include the visual of the instructor (audio+text) and a single-channel format (audio only). From a practical perspective, these two alternative modes of presenting lecture material, audio only and audio+text, are much easier to produce and easier to update from term to term than a full video lecture. Thus, in order to justify the use of video lectures that include the instructor in online learning there should be a clear benefit in terms of learning. Finally, given that one of the key differences between online and face-to-face lectures is the learning environment, which may directly impact attention, the environment the study is conducted in is critical. Thus, experiments should be conducted in an online environment rather than in a laboratory, where attention may be differentially impacted by dissimilarities in social factors and sources of distraction in these two environments. The present study aims to extend prior research by combining all of these important factors in a comprehensive manner to address this question of how video recordings of instructors in online lectures may impact attention, comprehension, and subjective preferences and beliefs about learning.

Across four experiments, we assess the potential impact of including video of an instructor on comprehension, mind wandering, and subjective evaluations of enjoyment, interest, and beliefs about how different lecture formats impact one's own learning. We examine the question of whether the instructor may act as a *seductive detail*, capturing attention and impairing comprehension or whether the instructor may facilitate comprehension by engaging attention and reducing internally generated thoughts or MW. We did not have a directional prediction of the impact of the instructor on self-reported MW. The exogenous or externally mediated capture of attention via the visual of the instructor may reduce mind wandering, helping learners focus attention on the lecture. On the other hand, the visual of the instructor may act primarily as a task irrelevant distractor, with little to no impact on MW. To assess MW we used a method of periodically probing thoughts throughout the lecture, which is commonly used to assess the proportion

of time spent mind wandering (Smallwood & Schooler, 2006). Learners may interpret the sense of engagement of attention with the visual of the instructor (whether this increases or reduces comprehension and MW) as an index of learning, leading to strong preferences for the video lecture with the instructor over other lecture formats.

In Experiment 1 and 2, different groups of online participants watched a lecture presented in one of several different formats (Experiment 1 – video with instructor, audio only, audio with text captions; Experiment 2 – video with instructor vs. audio with text). Throughout the lectures participants were probed about their MW. The impact of these different modalities on learning was assessed using a comprehension test following the lecture. In Experiments 3 and 4, we extended this research to explore learners explicit preferences across several different lecture formats (audio only, text only, audio+text, audio+instructor, audio+text+instructor) and their beliefs about how these lecture presentation formats impact their learning and likelihood of sticking with an online lecture.

Experiment 1

2.1 Methods

2.1.2 Participants. One hundred and sixty eight online participants, recruited from Amazon Mechanical Turk participated in Experiment 1. Participation was restricted to individuals located in the US. The sample consisted of 80 female participants, with a mean age of 33 years (range: 19-61). Participants were compensated with $3 and were debriefed via a feedback form following completion of the study. All procedures conformed to the University of Notre Dame's Institutional Review Board (IRB) guidelines and were approved by the IRB.

2.1.2 Design & Stimuli. A 3-factor between-subject design was employed, with random participant assignment to one of 3 lecture conditions: audio only (audio, N = 55, 19 females), audio with text (audio+text, N = 58, 21 females) and audio with a video of the instructor (audio+instructor, N = 55, 20 females). The same 24-minute segment of an introductory university-level Biology lecture on the topic of global problems of population growth was used in each condition and only the mode of presentation differed between conditions. The video was centered on the instructor, teaching in a lecture hall, and there were no slides, thus the dominant visual was of the instructor himself. The audio conditions had a static image of a lecture hall as the background with close captioned text superimposed at the bottom of the image for the audio+text condition.

2.1.3 Measures. Mind wandering was assessed using thought-probes at 13 different points throughout the lecture (see Figure 1 for task schematic). MW probes were presented as a beep tone, at which point participants indicated with a key press ('Y' for 'yes' or 'N' for 'no') if they were 'zoning out'

(mind wandering) just prior to the probe. The MW probes appeared at varying intervals with probe-to-probe lags ranging from about 80-165 seconds. These 13 probes appeared at the exact same time points in every lecture condition. The proportion of MW probes that participants reported they were MW ('zoning out') was used to estimate the proportion of time spent MW. For instance if a participant responded yes to 50% of the probes, they were estimated to be MW 50% of the time. Comprehension was assessed using a 16-question multiple-choice test, which measured memory for facts from the lecture, following the lecture (e.g., People wore wigs: a. to hide their own greying hair b. because they were unable to grow their own hair very long, c. to avoid lice, d. to demonstrate wealth). Comprehension was computed as proportion of correct responses.

2.1.4 Procedure. The study procedures and general purpose (i.e., to explore learning from online lectures) was communicated in written form to participants, after which participants provided electronic informed consent. Following consent, participants were provided more detail on the MW probes and specifically what 'zoning out' means and how to respond. They were then presented with the lecture which included the MW probes. Following the lecture and immediately preceding the comprehension test, participants indicated how interesting they found the lecture on a Likert scale (ranging from 1 or uninterested to 6 or very interested). Participants also answered individual difference questions assessing their level of alertness and everyday experiences of distractibility, however these two measures were not analyzed or discussed here. The entire procedure took about 40 minutes.

[Figure 1.]

2.2 Results

The impact of lecture format on lecture comprehension, proportion of MW, and subjective reports of interest in the lecture was examined using separate one-way ANOVAs with the between-subject factor of lecture format (audio, audio+text, audio+instructor). Lecture format significantly impacted lecture comprehension, $F(2, 166) = 3.14$, $p = 0.05$, $\eta^2 = .04$, such that comprehension was impaired in the audio+instructor format relative to the audio only format, $t(108) = 2.08$, $p = 0.04$, $\eta^2 = .04$, and the audio+text format, $t(112) = 2.04$, $p = 0.04$, $\eta^2 = .04$, while the audio and audio+text formats did not differ, $t(110) = 0.11$, $p = 0.91$ (see Figure 2.). Lecture format, however, did not significantly impact the proportion of time MW, $F(2, 166) = 0.06$, $p = 0.95$, $\eta^2 = .001$. Mind wandering, however, did show the typical negative association with comprehension, $r = -.32$, $p < 0.001$. Subjective level of interest in the lecture did not differ across lecture format, $F(2, 166) = 0.06$, $p = 0.95$, $\eta^2 = .001$

[Figure 2.]

2.3 Discussion

In Experiment 1 we found evidence that video lectures (audio+instructor) do not aid learning, in terms of comprehension and the ability to control attention (i.e., MW). Rather the video with the instructor impaired comprehension, relative to presenting the same lecture in an audio only format or audio with text format. This particular pattern of results (i.e., decrease in comprehension without an associated increase in MW) suggests that the video of the instructor is not impacting comprehension via increasing the likelihood of mind wandering. One potential interpretation of this effect is to view the instructor as a kind of seductive detail. Critically, paying attention to the instructor in and of itself could arguably impair processing of the relevant lecture content but may not be identified by participants as MW or 'zoning out', because they are paying attention to the instructor. In this sense, the cost in comprehension is not due to an increase in internal distraction as in MW, but rather may be due to external distraction as attention is captured by the instructor's visual features (Langton et al., 2008) and/or movements (Abrams & Christ, 2003). Lastly, participants did not report finding the different lecture formats differentially interesting.

3. Experiment 2

Experiment 2 replicates and extends the results of Experiment 1. It would be beneficial to replicate this effect with a different instructor in order to verify the effects from Experiment 1 were not unique to the particular instructor or lecture style. In addition, many online lectures are actually much shorter than the lecture used in Experiment 1, thus it is possible that the negative impact of the instructor might only emerge in lectures that are longer (the lecture in Experiment 1 was about 24 mins) and not those that are relatively brief. As such, it would be beneficial to determine whether the cost (and lack of benefit) generalize to a shorter lecture. We also extend our exploration of MW in Experiment 2 by using a different MW probe that affords a finer grained analysis of MW. Specifically, there is good evidence that people may MW intentionally or unintentionally (Seli, Carriere, & Smilek, 2015; Seli, Risko, & Smilek, 2016; Seli, Risko, Smilek, & Schacter, 2016) and participants in Experiment 1 may have only been characterizing the experience of spontaneous or unintentional MW as 'zoning out'. There are instances, however, when we *choose* to think about something other than the task at hand. For example, while watching a participant might remember that s/he has run out of milk, they realize they are no longer focusing on the lecture, but may choose to prioritize thinking about their to-do list or other grocery items they must remember to pick-up, or may simply allow their mind to wander more generally, off-task. These are still instances of MW (internally directed thought and disengagement from external task), but arguably may not be characterized by the same phenomenal aspects of feeling 'zoned out' or catching oneself in a spontaneous MW episode. Thus in Experiment 2 we explained the distinction between these two forms of MW and used MW probes

that ask participants to report on the kind of MW they were experiencing, intentional or unintentional. Finally, we were interested in the degree to which learners may be aware of their learning while watching these lectures and asked them to provide a judgment of learning (JOL) in the form of an estimate of the grade they expected on the test prior to the comprehension test.

3.1 Methods

3.1.1 Participants. Two hundred and six participants, recruited from Amazon Mechanical Turk, participated in Experiment 2. Participants were all located in either US or Canada. The sample consisted of 100 female participants, with a mean age of 34 years (range: 20-68). Sample size was determined in order to achieve 0.80 power to detect a medium effect size of $\eta^2 = 0.05$ with two-tailed, one-way omnibus ANOVAs, with an alpha criteria of .05. Participants were compensated with $3 and were debriefed via a feedback form following completion of the study. All procedures conformed to the guidelines of and were approved by the University of Waterloo's Research Ethics Board.

3.1.2 Design & Stimuli. A 2-factor between-subject design was employed, with random participant assignment to one of 2 lecture conditions: audio only (audio, N = 106, 47 females) or video of the instructor (audio+instructor, N = 100, 53 females). A different, and much shorter (6.5 mins), introductory History lecture on the topic of early modern England politics, religion and society, taught by a different instructor was used in Experiment 2. The same 6.5-minute lecture was used in each condition and only the mode of presentation differed between conditions. In the audio with instructor condition, the instructor was the prime focus of the video and there were no slides. In the audio condition, participants only saw a central image of a speaker and there was no image of a classroom as in Experiment 1.

3.1.3 Measures. Mind wandering was measured using thought-probes at 4 different points throughout the lecture. Experiment 2 employed a slightly different method of probing MW than that used in Experiment 1, which enables a way of distinguishing between different forms of MW. MW probes appeared as visual probes, asking participants if they were (1) fully focused and not mind wandering, (2) intentionally mind wandering or (3) unintentionally mind wandering. These two different forms of mind wandering were explained to participants in the instructions prior to the video. Intentional mind wandering was defined as: "Thoughts which are DELIBERATELY experienced that are not related to the material presented. For example, consciously thinking about what you will be making for dinner." Unintentional mind wandering was defined as: "Thoughts which are SPONTANEOUSLY experienced that are not related to the material presented. For example, a random thought about your friend coming to mind." Participants responded by using their mouse to select their attentional state in the moments just prior to the probe. Following the lecture, participants were asked to provide a judgement of learning (JOL), asking

"what percentage of the test questions do you think you will answer correctly ?" They also indicated how interesting they found the lecture using the same Likert scale from Experiment 1. Finally, comprehension was assessed using an 8-question multiple-choice test, that measured memory for facts from the lecture.

3.1.4 Procedure. The study procedure was the same as in Experiment 1, with the addition of the JOL question, which immediately followed the video and preceded the comprehension test. The procedure took about 15 minutes.

3.2 Results

The effect of lecture format on comprehension, JOLs, interest, and MW were examined using separate one-way ANOVAs, with a between-subject factor of lecture format (audio vs. video). Comprehension in the audio+instructor format was lower than in the audio only format, but this did not reach significance, $F(1, 205) = 2.93$, $p = 0.09$, $\eta^2 = .02$. Participants' JOLs were lower in the audio+instructor format, $F(1, 205) = 4.07$, $p = 0.05$, $\eta^2 = .02$, while subjective reports of interest did not differ between these two lecture formats, $F(1, 205) = 0.04$, $p = 0.84$, $\eta^2 < 0.001$ (see Figure 3). Overall MW, intentional MW, and unintentional MW did not significantly differ between these two lecture formats, Fs < 0.20, $p > 0.65$ (see Figure 4).

[Figure 3.]

[Figure 4.]

3.2.1 Combined Experiment 1 and 2 Analysis

We conducted a combined analysis of comprehension and MW by collapsing across the two experiments. We only included the audio and audio+instructor conditions (N=316) as there was no audio+text condition in Experiment 2. Before submitting the data to one-way ANOVAs, both the test scores and MW score were z-score standardized by experiment. Analysis of the standardized scores revealed a significant comprehension cost for the lecture containing the instructor, audio+instructor, relative to the audio only lecture, $F(1, 315) = 6.80$, $p = 0.01$, $\eta^2 = .02$. Again, there was no significant difference in MW between audio+instructor and audio only, $F(1, 315) = 0.09$, $p = 0.76$, $\eta^2 < .001$.

3.3 Discussion

The results of Experiment 2 further confirm that there is not a significant comprehension benefit associated with including visuals of the instructor relative to audio recordings. In fact, there appears to be a cost, which is confirmed in the analyses combining Experiments 1 and 2. In addition, as in Experiment 1,

there was no significant impact of lecture type on individuals' reports of MW. Individuals' JOL estimates closely mirrored actual test performance, being significantly higher for the audio condition, while subjective reports of interest in the lecture did not differ across these two modalities. As noted previously, this particular pattern of results is consistent with a seductive details mechanism. In particular, the instructor serves as a seductive detail that draws attention. Because paying attention to the instructor is likely not phenomenologically equivalent to completely disengaging from the lecture, this is not interpreted by the participant as mind wandering. Thus, including the instructor impairs comprehension with no appreciable effect on self-reported mind wandering.

4. Experiment 3

Experiment 1 and 2 provide evidence of what we refer to as the *Instructor Presence Effect*, whereby visuals of an instructor in an online video lecture impair learning. In Experiments 3 and 4 we sought to assess the potential influence of instructor presence on individuals' subjective evaluations of the lecture in a context where they are exposed to multiple types of lectures. In Experiments 1 and 2 participants' subjective evaluation of their learning (i.e., their JOLs) and interest were generated in what is referred to as *single* evaluation mode (Hsee & Zhang, 2010). This kind of comparison can underestimate the differences individuals perceive between different options as evidenced by subjective evaluations made in what is called *joint* evaluation mode. In the latter, individuals are provided with all of the relevant options and asked to make an evaluation. Participants' subjective evaluations in joint evaluation mode provide insight into their beliefs about how various options (i.e., lecture formats) compare against one another (e.g., in their learning, interest), as this type evaluation provides participants with relevant reference information (Hsee & Zhang, 2010). Thus, in Experiment 3 participants were presented with 4 different lecture modalities (i.e., audio only, text only, audio+text, and audio+instructor) and were asked to provide subjective ratings about their preferences (e.g., degree of enjoyment, which lecture they would choose to watch), beliefs about their learning (e.g., the extent to which the different lecture formats impact their learning and whether they believe they would drop lectures if presented in these different formats), and JOLs. This allows us to assess the underlying subjective evaluations that may motivate or guide students learning behavior in online courses. Provided the evidence in Experiments 1 and 2 that the presence of an instructor impairs learning, understanding individuals' subjective evaluations takes on added importance as these evaluations likely drive learning decisions (i.e., what type of lecture to view).

4.1 Methods

4.1.1 Participants. Eighty-eight participants, recruited from Amazon Mechanical Turk, located in either the US or Canada, participated in Experiment 3. The sample consisted of 42 female participants, with a mean age of 33 years (range: 19-60). Sample size was estimated in order to achieve 0.80 power to detect a medium effect size of $\eta^2 = 0.04$ with two-tailed, repeated-measures ANOVA, with an alpha criteria of .05. Participants were compensated $4 and were debriefed via a feedback form following completion of the study. All procedures conformed to the guidelines of and were approved by the University of Waterloo's Research Ethics Board.

4.1.2 Design & Stimuli. The same 6.5 minute lecture used in Experiment 2 was divided roughly into 4 segments of equal duration (~1.5 mins each) and 4 different lecture formats were generated for each segment: audio only, text only, audio+text, or audio+instructor. The order of lecture format segment was counter-balanced across participants, producing 24 different combinations of format order. This ensures that systematic differences in outcome variables across lecture modality cannot be attributed to variance in lecture content/difficulty across the lecture or time on task. In between lecture segments, participants were presented with an instruction screen that signified the end of a segment and indicated the format in which the next segment would be presented. A 4-factor (audio, text, audio+text, audio+instructor) within-subject design was employed, with random participant assignment to one of 24 lecture conditions (see Figure 5).

[Figure 5.]

4.1.3 Measures. Following viewing all four lecture segments, participants were asked to rate their experience of and preference for each of the different lecture formats using a Likert scale, ranging from 1-5 (ranging from 'none/not at all' to 'a lot/very much'). Participants were asked to rate the extent to which they found each lecture format: enjoyable, interesting, and the degree of effort required to maintain attention on the lecture (attentional effort). Participants were also asked about their beliefs about their learning and to make predictions about their learning outcomes. For each lecture format, participants used the same Likert scale to report the degree to which they believed that lecture format facilitated their learning, how likely they would be to drop a course presented mainly in that lecture format, and how likely they would be to stick with a lecture to completion if presented in that lecture format. Participants were also asked to select the one lecture format that they believe they learn best from, then they were asked which they enjoyed the most, and then were asked which they most prefer. Finally, participants to provided a judgement of learning (JOL) for each lecture format by estimating the percentage of comprehension questions they expected to answer accurately.

Comprehension was assessed using an 8-question multiple-choice test following the lecture, where 2 questions probed content from each segment of the lecture. Comprehension was not the focus of the present investigation and questions were included to encourage participants to focus on the material. The inclusion of comprehension tests also enables the collection of meaningful JOLs. MW was not measured in this experiment.

4.1.4 Procedure. Just as in Experiment 1 and 2, participants were first informed of the study details and then they watched the lecture segments. The lecture was followed by the subjective perceptions survey, JOL estimates, and then the comprehension test. The procedure took about 30 minutes.

4.2 Results

4.2.1 Preferences and Beliefs about Lecture Format

The question of whether learners may hold preferences or beliefs about how different lecture formats in general and how lectures with visuals of the instructor in particular impact their learning was explored by submitting the subjective report data to separate repeated-measures ANOVAs with a within-subject factor of lecture format (audio, text, audio+text, audio+instructor). Simple effects are followed-up with planned pair-wise comparisons between the audio+instructor format and the other 3 formats. There was a main effect of enjoyment, $F(3, 261) = 35.65$, $p < 0.001$, $\eta^2 = .29$, with participants reporting significantly higher enjoyment for audio+instructor format relative to audio, $t(87) = 6.65$, $p < 0.001$, $\eta^2 = .15$, and text, $t(87) = 7.23$, $p < 0.001$, $\eta^2 = .21$, but not audio+text, $t(87) = 0.53$, $p = 0.56$, $\eta^2 = .001$. (see Figure 6). Participants also showed a significant difference in their ratings of interest for the different lecture formats, $F(3, 261) = 35.45$, $p < 0.001$, $\eta^2 = .26$, with significantly higher ratings for the audio+instructor format relative to audio alone, $t(87) = 4.88$, $p < 0.001$, $\eta^2 = .18$, and text alone, $t(87) = 7.23$, $p < 0.001$, $\eta^2 = .18$, but not audio+text, $t(87) = 0.51$, $p = 0.61$, $\eta^2 < .001$. Participants' self-reported attentional effort also differed across lecture format, $F(3, 261) = 16.90$, $p < 0.001$, $\eta^2 = .16$, with participants reporting significantly lower attentional effort in the audio+instructor format relative to audio format, $t(87) = 4.56$, $p < 0.001$, $\eta^2 = .05$, and the text format, $t(87) = 4.94$, $p < 0.001$, $\eta^2 = .08$, but not the audio+text format, $t(87) = 1.56$, $p = 0.12$, $\eta^2 < 0.001$.

Participants also believed the lecture formats had significantly different impacts on their learning, $F(3, 261) = 37.54$, $p < 0.001$, $\eta^2 = .30$. Specifically, they believed that the audio+instructor format leads to better learning relative to the audio format, $t(87) = 6.75$, $p < 0.001$, $\eta^2 = .20$, and the text format, $t(87) = 7.30$, $p < 0.001$, $\eta^2 = .25$, but not the audio+text format, $t(87) = 1.34$, $p = 0.18$, $\eta^2 < 0.001$. Further, participants' predictions of whether they would *drop a course* presented primarily in one of these lecture formats differ significantly across lecture format, $F(3, 261) = 35.67$, $p < 0.001$, $\eta^2 = .29$, with a significantly

lower likelihood of dropping a course presented in the audio+instructor format relative to the audio format, $t(87) = 6.62$, $p < 0.001$, $\eta^2 = .15$, and the text format, $t(87) = 7.61$, $p < 0.001$, $\eta^2 = .22$, but not the audio+text format, $t(87) = 1.60$, $p = 0.11$, $\eta^2 = .001$. Participants' beliefs about the *likelihood of sticking with a lecture to completion* significantly differs across lecture format, $F(3, 261) = 32.03$, $p < 0.001$, $\eta^2 = .27$, with significantly lower likelihoods of dropping in the audio+instructor format relative to the audio format, $t(87) = 7.71$, $p < 0.001$, $\eta^2 = .20$, and the text format, $t(87) = 7.28$, $p < 0.001$, $\eta^2 = .23$, but not the audio+text format, $t(87) = 1.45$, $p = 0.15$, $\eta^2 = .01$.

[Figure 6.]

Participants' JOLs did show a significant main effect of lecture format, $F(3, 261) = 2.81$, $p = 0.04$, $\eta^2 = .03$ (Figure 6). However, none of the planned pair-wise comparisons were significant, all $t < 1.7$, $p>0.09$. Instead, follow-up pair-wise comparisons indicated that the main effect was driven by participants' belief that they would score higher on the portion of the test that contained content from the text only section, relative to the audio only condition, $t(87) = 2.86$, $p = 0.005$, $\eta^2 = .02$. JOLs for text only were marginally higher relative to audio+text, $t(87) = 1.85$, $p = 0.07$, $\eta^2 = .01$, and audio+instructor, $t(87) = 1.73$, $p = 0.09$, $\eta^2 = .01$. No other pair-wise comparison was significant, all $p>0.25$.

When we asked participants to choose the one lecture format that they believe they learn the best in, enjoy the most, and prefer overall (see Table 1), descriptive data shows that about half of participants believe they learn best in the audio+instructor format, while the other half believe they learn best in the audio+text format. Most people appear to enjoy the audio+instructor format, but were roughly divided across the audio+instructor or the audio+text format in terms of beliefs about learning and preferences.

[Table 1.]

4.2.2 Comprehension

There were no systematic effect of lecture format on comprehension, $F(3, 261) = 1.15$, $p = 0.33$, $\eta^2 = .01$.

4.3 Discussion

The results of Experiment 3 suggest that learners prefer and believe that their learning is facilitated by both multi-channel formats - the audio+text and the audio+instructor formats. Thus, they may be more likely to choose to watch lecture material that is multi-channel and/or includes visuals of the instructor. The JOL results suggest that learners estimate higher learning in the text only condition, relative to the audio only condition, but there was no differences between the instructor format and any other condition. Note that this contrasts with the results in Experiment 2 wherein individuals' JOLs were lower in the

audio+instructor condition. This likely reflects the different evaluation modes used (i.e., single vs. joint evaluation) – see General Discussion.

5. Experiment 4

In Experiment 3, the audio+text and audio+instructor conditions yielded similar subjective evaluations. This likely reflects the fact that both include two "channels" of information, relative to the one channel present in the audio and text only conditions. Interestingly, in the case of the audio+instructor condition, the added visual channel arguably adds no "new" information (in fact according to Experiments 1 and 2 it impairs comprehension) whereas adding text could, in theory, be beneficial. For instance, individuals learning a different language benefit from the use of text captioning and presentation formats that combine media, such as text with visual information (Brett, 1997; Garza, 1991), as providing text may enable language learners to catch up if the spoken audio becomes too fast for them to follow. To provide further evidence that adding the instructor influences subjective evaluations (e.g., increases preference, beliefs about learning), we examined the impact of adding the instructor to the audio+text format in Experiment 4. Thus, the audio+text and audio+text+instructor comparison differs only in terms of the presence of the instructor, which makes the comparison more controlled.

5.1 Methods

5.1.1 Participants. One hundred and nine participants, recruited from Amazon Mechanical Turk, located in either US or Canada, participated in Experiment 4. The sample consisted of 36 female participants, with a mean age of 34 years (range: 21-63). Sample size was estimated in order to achieve 0.80 power to detect a medium effect size of $\eta^2 = .025$ (found in Experiment 3) with two-tailed, repeated-measures ANOVA, with an alpha criteria of .05. Participants were compensated with $4 and were debriefed via a feedback form following completion of the study. All procedures conformed to the guidelines of and were approved by the University of Waterloo's Research Ethics Board.

5.1.2 Design & Stimuli. The exact same 6.5 minute lecture divided roughly into 4 segments of equal duration (~1.5 mins each) used in Experiment 3 was used here. The only difference was the addition of text to the audio+instructor condition (now called audio+text+instructor). The 4 different lecture formats were as follows: audio only, text only, audio+text, or audio+text+instructor (Figure 5). The order of lecture format segment was counter-balanced across participants, producing 24 different combinations of format order in the exact same manner as Experiment 3.

5.1.3 Measures & Procedure. The exact same subjective reports, JOL estimates and comprehension test from Experiment 3 were used. The procedure took about 30 minutes.

5.2 Results

5.2.1 Preferences and Beliefs about Lecture Format

Participants' subjective report data was submitted to separate repeated-measures ANOVAs, with a within-subject factor of lecture format (audio, text, audio+text, audio+text+instructor). Simple effects are followed-up with planned pair-wise comparisons between the audio+text+instructor format and the other 3 formats. Participants show evidence of a clear preference for the lecture containing the visual of the instructor, as enjoyment significantly differed across lecture format, $F(3, 324) = 89.44$, $p < 0.001$, $\eta^2 = .45$. Participants report significantly higher enjoyment for audio+text+instructor format relative to audio only, $t(108) = 10.73$, $p < 0.001$, $\eta^2 = .30$, text, $t(108) = 12.52$, $p < 0.001$, $\eta^2 = .39$, as well as audio+text, $t(108) = 5.35$, $p < 0.001$, $\eta^2 = .08$. (see Figure 7). Participants also showed a significant difference in their ratings of interest for the different lecture segments, $F(3, 324) = 64.17$, $p < 0.001$, $\eta^2 = .37$, with significantly higher interest reported for the audio+text+instructor format relative to audio alone, $t(108) = 7.78$, $p < 0.001$, $\eta^2 = .18$, text alone, $t(108) = 10.75$, $p < 0.001$, $\eta^2 = .32$, as well as the audio+text format, $t(108) = 3.80$, $p < 0.001$, $\eta^2 = .08$. Participants' subjective sense of attentional effort also differed across lecture format, $F(3, 324) = 40.27$, $p < 0.001$, $\eta^2 = .27$, with participants reporting significantly lower attentional effort in the audio+text+instructor format relative to audio format, $t(108) = 6.94$, $p < 0.001$, $\eta^2 = .10$, the text format, $t(108) = 8.19$, $p < 0.001$, $\eta^2 = .18$, and the audio+text format, $t(108) = 2.22$, $p = 0.03$, $\eta^2 = .01$.

Participants report that these lecture formats do not equally support their learning, $F(3, 324) = 79.91$, $p < 0.001$, $\eta^2 = .43$, as they believe that the audio+text+instructor format leads to better learning relative to the audio format, $t(108) = 12.34$, $p < 0.001$, $\eta^2 = .38$, the text format, $t(108) = 11.47$, $p < 0.001$, $\eta^2 = .38$, and the audio+text format, $t(108) = 4.95$, $p < 0.001$, $\eta^2 = .08$ (see Figure 7). Participants' predictions of whether they would *drop a course* presented primarily in one of these lecture formats also differ significantly across lecture format, $F(3, 324) = 65.01$, $p < 0.001$, $\eta^2 = .38$, as they report significantly lower likelihood of dropping a course presented in the audio+text+instructor format relative to the audio format, $t(108) = 9.68$, $p < 0.001$, $\eta^2 = .29$, the text format, $t(108) = 10.20$, $p < 0.001$, $\eta^2 = .34$, and audio+text format, $t(108) = 5.35$, $p < 0.001$, $\eta^2 = .08$. Participants' beliefs about the likelihood of *sticking with a lecture to completion* significantly differs across lecture format, $F(3, 324) = 62.16$, $p < 0.001$, $\eta^2 = .37$, as they report significantly higher likelihood of watching a lecture to completion in the audio+text+instructor format relative to the audio format, $t(108) = 8.42$, $p < 0.001$, $\eta^2 = .23$, the text format, $t(108) = 11.20$, $p < 0.001$, $\eta^2 = .30$ and audio+text format, $t(108) = 5.57$, $p < 0.001$, $\eta^2 = .09$.

Participants' JOLs differed significantly with lecture format, $F(3, 324) = 25.90$, $p < 0.001$, $\eta^2 = .19$ (Figure 7). JOLs were significantly higher in the audio+text+instructor format relative to the audio only

format, $t(108) = 5.94$, $p < 0.001$, $\eta^2 = .05$, the text format, $t(108) = 6.92$, $p < 0.001$, $\eta^2 = .08$, and the audio+text format, $t(108) = 2.45$, $p = 0.01$, $\eta^2 = .01$.

[Figure 7.]

When participants are asked to choose the one lecture format that they believe they learn best in, enjoy the most, and prefer overall, descriptive data shows that more than 70% of people choose the audio+text+instructor format (see Table 1.)

5.2.2 Comprehension

There was a marginal effect of lecture format on comprehension, $F(3, 324) = 2.33$, $p = 0.07$, $\eta^2 = .02$. Planned pair-wise comparisons indicated that comprehension was significantly lower in the audio+text+instructor format ($M = 0.55$, $SD = 0.38$) relative to the audio only format ($M = 0.64$, $SD = 0.37$), $t(108) = 2.05$, $p = 0.04$, $\eta^2 = .01$ and the audio+text format ($M = 0.66$, $SD = 0.38$), $t(108) = 2.56$, $p = 0.01$, $\eta^2 = .02$, but only marginally to the text format ($M = 0.63$, $SD = 0.38$), $t(108) = 1.69$, $p = 0.09$, $\eta^2 = .01$.

5.2.3 Gender Effects

The random sample collected for this experiment contained notably fewer females. To ensure that these results were not biased by gender difference point-biserial correlations were conducted to see if gender was predictive of any of our outcome measures. There are no significant correlations, except for one small relationship, revealing higher reports of attentional effort associated with females, $r = .2$, $p = 0.01$. To ensure that the pattern of results for subjective evaluations of attentional effort did not significantly differ by gender a mixed ANOVA was conducted, with lecture format as a within-subject factor and gender as a between-subject factor. While females' reports of attentional effort generally tended to be higher than males, $F(1, 107) = 6.61$, $p = 0.01$, $\eta^2 = .06$, there is no significant interaction between gender and lecture format, $F(1, 107) = 0.006$, $p = 0.94$, $\eta^2 = .00$, thus males and females show the same overall effect of lecture format on attentional effort.

5.3 Discussion

By adding text to the audio+instructor condition in Experiment 4 we equated the audio+text and audio+text+instructor condition in all ways except for the addition of the instructor. This manipulation reveals that learners have a clear preference for and belief that they learn best when video lectures contain visuals of the instructor. Further, they believe lecture material containing the instructor benefits their learning and they predict that they are more likely to take a course and stick with a lecture to completion if

the instructor is present. JOLs reflect a similar pattern, as individuals estimate/predict their learning and performance to be highest when video lectures contain visuals of the instructor. There is, however, a marginal effect on comprehension, suggesting comprehension may be reduced in the audio+text+instructor condition, although this result should be interpreted with caution, as this measure contained few comprehension questions per lecture.

A potential limitation to this study is the unequal distribution of females/males in the random sample and evidence of a small relationship between gender and attentional effort. While females, overall, may provide slightly higher subjective ratings of attentional effort, the pattern of subjective ratings of attentional effort across the different lecture formats did not differ between males and females.

6. General Discussion

The present set of experiments provide evidence that adding visuals of an instructor to video lectures has a negative impact on learning relative to other formats (audio, text, audio+text), while most learners appear to prefer and believe that this format best serves their learning. The results of Experiments 1 and 2 reveal, using two different lectures of different durations (~24 mins and ~6.5 mins), that comprehension is not benefitted by adding visuals of the instructor in video lectures. Rather, visuals of the instructor (audio+instructor) negatively impacts comprehension, relative to audio only and audio+text formats. The results do not support the notion that the effect is driven by an increase in internally generated distraction or an inability to keep attention focused on the task at hand, as we find no differences in MW across these lecture formats. Rather, this comprehension cost may arise from the externally generated visual distraction of the instructor. Given that visual features, such as faces (Langton et al., 2008) and movement (Abrams & Christ, 2003) capture attention in a rather automatic way, the salient visual features of the instructor may act to capture attention, consuming capacity-limited resources that otherwise would be used for processing the conceptual content of the lecture.

This negative impact of salient but irrelevant information has on learning could be considered a *seductive detail* (Harp & Mayer, 1998; Mayer, Griffith, Jurkowitz, & Rothman, 2008; Mayer et al., 2001; Sung & Mayer, 2012). In *the laboratory* Mayer and colleagues have shown that the use of entertaining details or *images* that are not germane to the concept being taught in a lecture, may be enjoyed by learners, but can have a deleterious impact on learning (Harp & Mayer, 1998; Sung & Mayer, 2012). The novel contribution of the research presented here is that *visuals of the instructor in online video lectures* may actually act as a seductive detail, resulting in a cost to comprehension, which we refer to as the *instructor presence effect*. Learners may also interpret their attention being captured by the instructor as being engaged with the lecture, feeling an increased sense of ease in maintaining attention. They may interpret this as a sign that

they must be learning more/better. What they may not realize is that they are attending to the wrong information or missing relevant conceptual content.

Although visuals of the instructor in video lectures impairs comprehension, learners seem to like it. Consistent with prior work on seductive details and instructor-personalized learning materials (Gunawardena & Zittle, 1997; Mandernach, 2009), Experiments 3 and 4 provide evidence that adding visuals of the instructor boosts learners enjoyment and interest (i.e., increasing with addition of visual of instructor to audio in Experiment 3 and addition of visual of instructor to audio+text in Experiment 4). Learners report that they find lectures easier to attend to, and believe they learn better when the instructor is present. They also predict that they are more likely to choose courses are less likely to drop lectures in this format. Given the considerable control and freedom online learners have over how and when they learn, these subjective evaluations are likely to guide the decisions that learners make in determining the learning resources they use, impacting their learning outcomes. Thus, it is important to take these evaluations into consideration when deciding how to present lecture content.

Like all studies, ours has limitations that ought to be considered. For one, the proportion of time spent MW during lectures is typically around 40% (Hollis & Was, 2016; Risko et al., 2012), while the proportion of MW in our studies was closer to 20%. It is unclear what might underlie the lower proportion of MW in the present data, as the two different lectures differed in content, instructor, duration, and type of MW probe, but proportion of MW was about 20% for each. It is possible that our online participants under-reported their MW. Alternatively, they may have accurately reported MW, but were perhaps motivated to perform well on the task, which can lead to a reduction in MW (Seli, Wammes, Risko, & Smilek, 2015). Though it remains unclear why the proportion of MW was lower overall in these studies relative to previous work, this should not impact the main finding that comprehension is negatively impacted when visuals of an instructor are included in online lectures.

We also note that the JOL results are somewhat inconsistent across experiments, which may be due, in part, to differences in the context in which participants made this evaluation between Experiment 2 and the last two experiments. In Experiment 2, participants were exposed to just one lecture format and participants' JOLs were lower for the audio+instructor format, relative to audio only format, consistent with test performance. In Experiment 3, however, participants did not predict a difference between the audio only and audio+instructor formats, but expected slightly higher performance in the text only condition. In Experiment 4 participants JOLs were highest for the audio+text+instructor format. As noted prior to Experiment 3 one way Experiment 2 differed from 3 and 4 was in terms of the decision context in which participants were making this evaluation, single vs. joint evaluation (Hsee & Zhang, 2010). Experiment 3 and 4 provide participants with a reference by which to evaluate their learning across

different formats (Hsee & Zhang, 2010). In a real course, students make joint evaluations when deciding which materials to learn from and the design in Experiment 3 and 4 provides a more appropriate context for revealing learners' beliefs about the relative utility of different lecture formats. By similar reasoning, we may explain why Experiment 1 and Experiment 2 did not show a significant difference in interest between lecture formats, which diverges from the clear preference for lecture formats with visuals of the lecturer demonstrated in Experiments 3 and 4. This single evaluation in a between-subject design may simply not be sensitive enough to pick-up on differences in subjective interest-level as a function of lecture format.

An important consideration, which is not accounted for or tested in the present set of experiments, is the important role an instructor likely plays in fostering a sense of *social connection* and *motivating* students to learn over time, throughout the duration of a course. The lectures in these studies were single lectures, unlike most online courses, which include many lectures, which are watched over the duration of a course, typically extending over months. One of the major challenges in online courses, similar to face-to-face lectures, is motivating students to 'show-up', week after week. Students' decision to engage with online learning materials is likely based, in part, on a combination of how they feel (enjoyment and interest) while engaging with the learning materials, the degree to which they feel or believe they are benefiting from that learning material (sense of learning), and whether they think the learning material presented is going to help them actually learn (JOL) or achieve their goal (which may be to simply earn a certificate or grade). Given the challenge of engaging online learners and motivating them to show-up and persist through the entire duration of a video lecture (Kim et al., 2014), there is something to be said for including visuals of the instructor, despite a relative comprehension cost. If students are dropping lectures only 5 minutes in, as Kim et al's. (2014) study reveals, or not showing up at all, there is likely a cost to comprehension, which is not captured in the present set of experiments. Of course, this assumes that the presence of an instructor in the videos does lead to enhanced engagement and persistence outcomes.

Another potential limitation to studying learners in these single lectures is that motivation levels may differ for these online experiments relative to live classes and motivation may interact with these effects. For instance, students that are particularly motivated to learn (because of interest or desire to perform well in a course) may show less of a comprehension cost with the addition of visuals of the instructor. Although some students may be motivated by marks/points and a desire to perform well in a real course, others are likely not particularly motivated. Nonetheless, individual differences in learners' motivation, preferences, beliefs and the lecture materials they choose to engage with should be explored further in real live courses over the duration of a semester investigating. In addition, motivation effects may differ across different disciplines. The instructor may act as a greater distractor or seductive detail in lectures that are more visual (e.g., understanding the Krebs cycle or structures of a cell in a Biology lecture)

relative to those that are more narrative, such as History or English lectures. In the present study was limited to just two different disciplines, 2 lectures, 2 instructors, and 2 durations, thus, future work should explore the greater generalizability of these results across instructors and subject matter, as well as durations of lectures and courses.

There are also additional unmeasured individual difference factors which may influence these effects. For instance, the instructor presence effect and/or adding text to a lecture may differ depending on whether a student is: an English language learner; a novice or expert; a child or adult learner. Thus these results should be taken with some caution and should not be over generalized. While further research is required to determine the extent to which these results generalize across learner populations and disciplines, the present results provide some insight into a tension between liking and learning in the context of online education.

Looking towards future work, these results illustrate the importance of understanding both what students prefer and what they believe will help them learn, as well as what will in fact best support comprehension and learning. With this knowledge, we can better design online learning materials that are able to achieve both the goal of motivating students and the goal of presenting material in a way that reflects sound learning principles. For instance, there may be a more strategic way of including a video of the instructor to motivates and engages students without distracting them from core concepts, as Fiorella and Mayer's (2016) results suggest that visuals of the instructor or parts of the instructor can enhance learning. The instructor may attract attention and act as a social cue, but this may help or hurt learning, depending on how and when the instructor is presented. In one study, Kizilcec et al., (2015) tested strategic presentation of an instructor in a video by having the instructor appear at key points, but this did not aid learning and some reported it to be distracting. An alternative approach might be to present the instructor at the beginning and end of the lecture, providing learners with context and fostering a sense of social connection that may persist throughout the lecture due to the presence of the instructor's voice without an attention-capturing visual. Rather, the visual of the instructor would be included as a motivational cue and would be removed when learners need to focus attention on learning material.

Conclusion

We demonstrated an *instructor presence effect* in the present set of experiments. The effect is characterized by a cost to comprehension when visuals of an instructor are present in an online video lecture despite learners preferring and believing that this leads to the best learning. In Experiments 1 and 2, we found evidence that the visual presence of an instructor in a video lecture leads to a comprehension cost, which may be due to attention being captured by the visual features of the instructor rather than mind

wandering. Consequently, Experiments 3 and 4 revealed that subjective ratings of attention and learning are increased by the visual presence of the instructor in the video over audio alone and text alone (Experiment 3) and audio+text (Experiment 4). This clear preference may arise from the subjective sense of attention being captured and engaged by the instructor, while learners may not be aware of the content they are missing or processing in a more shallow fashion as a result. This research illustrates a clear gap between learning and liking, but also demonstrates the importance of designing online courses that work to close this gap. If we are to truly *design with the learner in mind*, we must recognize that online learners have considerable flexibility and freedom over their own learning, thus their preferences and beliefs influence how, what, and when they learn. At the same time, we must also be aware of potential cognitive and learning costs to said preferences and beliefs. This gap between learning and liking may be bridged by gaining a better understanding of, and bringing together, growing knowledge of cognitive and emotional influences on learning and liking in online learning environments

References

Abrams, R. A., & Christ, S. E. (2003). Motion onset captures attention. *Psychological Science*, *14*(5), 427–432.

American Association of Community Colleges. (2015). Distance Education Survey Results. Trends in eLearning: Tracking the impact of eLearning at community colleges. Retrieved from http://www.itcnetwork.org/resources/1439-itc-2015-distance-learning-survey-results.html

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*(1), 117–139.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System*, *25*(1), 39–53.

Chaiken, S., & Eagly, A. H. (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology*, *34*(4), 605.

Christoff, K., Gordon, A., Smallwood, J., Smith, R., & Schooler, J. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences*, *106*(21), 8719–8724.

Dixon, M. L., Ruppel, J., Pratt, J., & De Rosa, E. (2009). Learning to ignore: Acquisition of sustained attentional suppression. *Psychonomic Bulletin & Review*, *16*(2), 418–423.

Farley, J., Risko, E., & Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*, 619.

Garza, T. J. (1991). Evaluating the Use of Captioned Video Materials in Advanced Foreign Language Learning. *Foreign Language Annals*, *24*(3), 239–258.

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*(3), 8–26.

Harp, S. F., & Mayer, R. E. (1998). How seductive details do their damage: A theory of cognitive interest in science learning. *Journal of Educational Psychology*, *90*(3), 414.

Heidig, S., Müller, J., & Reichelt, M. (2015). Emotional design in multimedia learning: Differentiation on relevant design features and their effects on emotions and learning. *Computers in Human Behavior*, *44*, 81–95.

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience*, *18*(4), 604–613.

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104–112.

Hsee, C. K., & Zhang, J. (2010). General evaluability theory. *Perspectives on Psychological Science*, *5*(4), 343–355.

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks inonline lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 31–40).

Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). The instructor's face in video instruction: Evidence from two large-scale field studies. *Journal of Educational Psychology*, *107*(3), 724.

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 32nd annual ACM conference on Human factors in computing systems* (pp. 2095–2102). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2557207

Kopp, K., & D'Mello, S. (2016). The Impact of Modality on Mind Wandering during Comprehension. *Applied Cognitive Psychology*, *30*(1), 29–40.

Langton, S. R., & Bruce, V. (1999). Reflexive visual orienting in response to the social attention of others. *Visual Cognition*, *6*(5), 541–567.

Langton, S. R., Law, A. S., Burton, A. M., & Schweinberger, S. R. (2008). Attention capture by faces. *Cognition*, *107*(1), 330–342.

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59.

Mandernach, B. J. (2009). Effect of instructor-personalized multimedia in the online classroom. *The International Review of Research in Open and Distributed Learning*, *10*(3). Retrieved from http://www.irrodl.org/index.php/irrodl/article/viewArticle/606

Martel, C. (2015). *Online and distance education capacity of Canadian universities*. Retrieved from http://www.tonybates.ca/2016/03/23/a-national-survey-of-university-online-and-distance-learning-in-canada/

Mayer, R. E., Griffith, E., Jurkowitz, I. T., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, *14*(4), 329.

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*(1), 187.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: Role of speaker's voice. *Journal of Educational Psychology*, *95*(2), 419.

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*(5), 857–863.

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*(2), 177–213.

Mousavi, S. Y., Low, R., & Sweller, J. (1995). Reducing cognitive load by mixing auditory and visual presentation modes. *Journal of Educational Psychology*, *87*(2), 319.

Otondo, R. F., Van Scotter, J. R., Allen, D. G., & Palvia, P. (2008). The complexity of richness: Media, message, and communication outcomes. *Information & Management*, *45*(1), 21–30.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive load theory and instructional design: Recent developments. *Educational Psychologist*, *38*(1), 1–4.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*(2), 234–242.

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*, *68*, 275–283.

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, *73*(2), 291–296.

Seli, P., Carriere, J. S., & Smilek, D. (2015). Not all mind wandering is created equal: Dissociating deliberate from spontaneous mind wandering. *Psychological Research*, *79*(5), 750–758.

Seli, P., Risko, E. F., & Smilek, D. (2016). On the necessity of distinguishing between unintentional and intentional mind wandering. *Psychological Science*, *27*(5), 685–691.

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences*, *20*(8), 605–617.

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2015). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review*, 1–8.

Smallwood, J. (2013). Distinguishing how from why the mind wanders: A process–occurrence framework for self-generated mental activity. *Psychological Bulletin*, *139*(3), 519–535.

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946.

Sung, E., & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*, *28*(5), 1618–1625.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review*, *11*(1), 65–70.

Theeuwes, J., & Burger, R. (1998). Attentional control during visual search: The effect of irrelevant singletons. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(5), 1342–1353.

Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, *13*(6), 657–665.

Acknowledgements

This research was supported by a Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Discovery Grant, funding from the Canada Research Chairs program, and an Early Researcher Award from the Province of Ontario to EFR and a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant to DS.

Tables

				Lecture Format		
Experiment	Question	audio	text	audio+ text	audio+ instructor	audio+text+ instructor
Exp3	Learn Best	0	3	47	50	-
	Enjoy Most	0	1	28	71	-
	Prefer Most	0	2	42	56	-
Exp4	Learn Best	1	8	19	-	72
	Enjoy Most	2	4	20	-	74
	Prefer Most	1	6	21	-	72

Table 1. Experiment 3 & 4: Proportion of participants who selected each format as the format they learn best from, enjoy the most, and simply prefer the most.

Figures

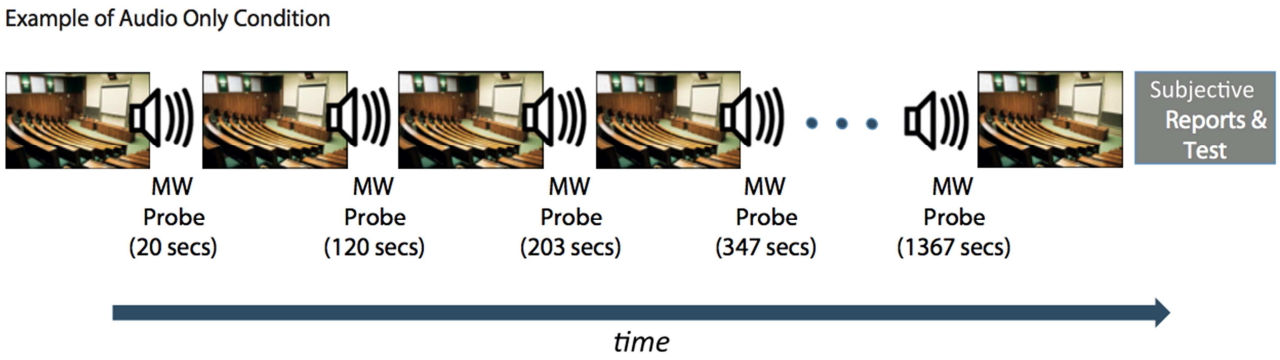

Figure 1. Experiment 1 Task Schematic: Task procedure with MW probes (audio beeps) throughout lecture, followed by subjective reports and the comprehension test. Example is of audio only condition. Audio with text (audio+text) condition was same image with text subtitles along the bottom quarter of the screen. The Audio with instructor (audio+instructor) condition was a video of the instructor teaching at the front of a lecture hall.

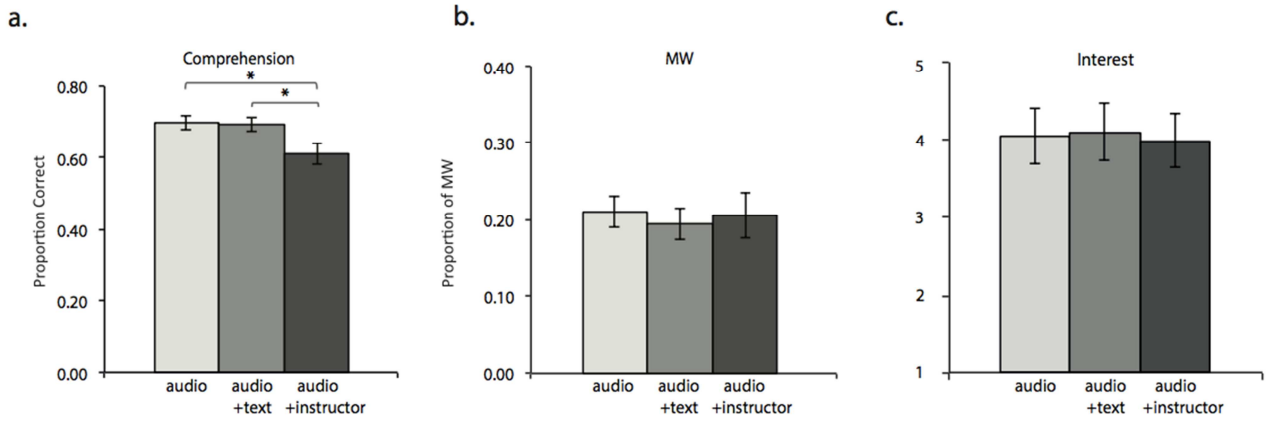

Figure 2. Experiment 1: The effect of lecture format on a.) comprehension, and b.) mind wandering. There is a significant main effect of lecture format on comprehension, but no effect of MW. c.) Participants' subjective ratings of interest in the lecture (5-point Likert scale) did not differ across lecture format. Error bars reflect standard error of the mean (SEM). * indicates $p < 0.05$.

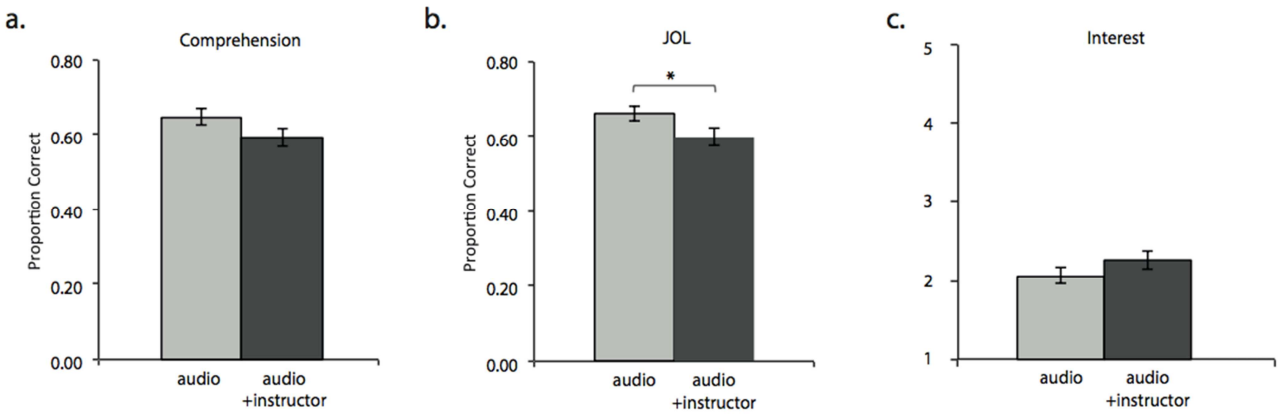

Figure 3. Experiment 2: The effect of lecture format on a.) Comprehension and b.) Participants' Judgments of learning (JOL). There is a small but not quite significant effect of lecture format on comprehension and a significant effect of lecture format on JOLs. c.) Participants' subjective ratings of interest in the lecture (5-point Likert scale) did not differ across lecture format. Error bars reflect standard error of the mean (SEM). * indicates $p < 0.05$.

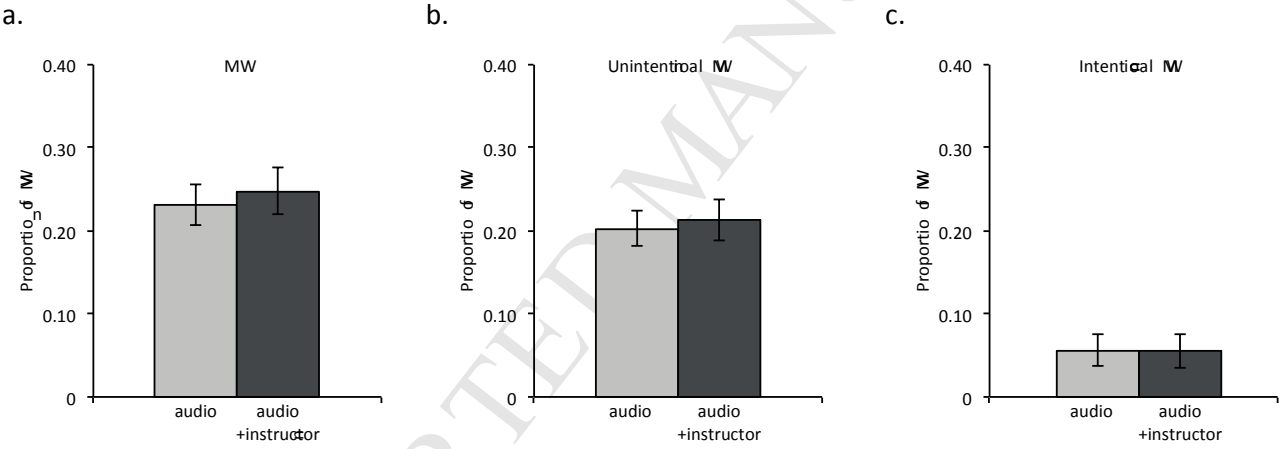

Figure 4. Experiment 2: The effect of lecture format on a.) Mind Wandering in general, b.) Unintentional MW, c.) Intentional MW. There is no significant main effect of lecture format on MW. Error bars reflect standard error of the mean (SEM).

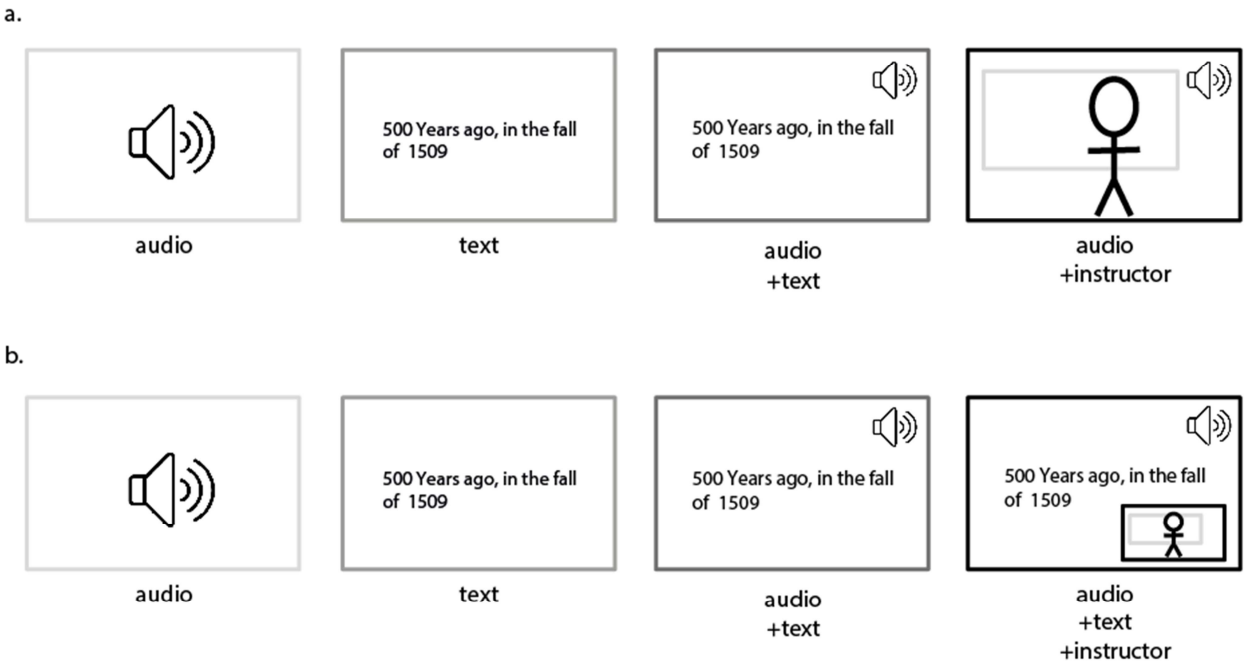

Figure 5. Experiment 3 and Experiment 4 Stimuli. Both experiments use the same 6.5 minute history lecture. a.) Experiment 3 and b.) Experiment 4 differ in one lecture modality, where text was added to the audio with instructor video in Experiment 4.

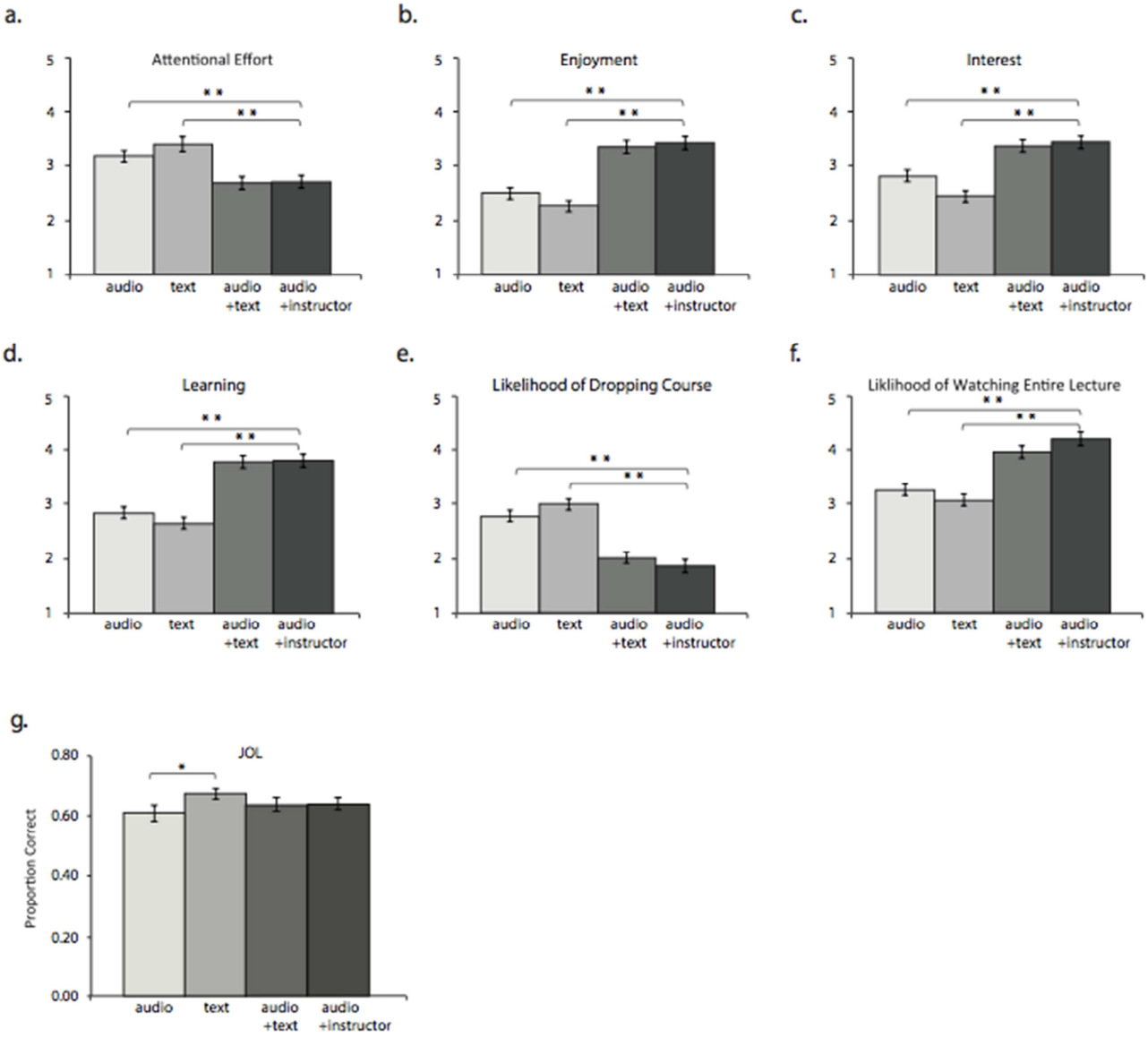

Figure 6. Experiment 3: Subjective ratings (Likert Scale, 1-5, ranging from a lot/very to very little). a.) amount of effort maintaining attention across different formats, b.) degree of enjoyment watching lecture across different formats, c.) level of interest watching lecture across different formats, d.) how much learning is facilitated across different formats, e.) predicted likelihood of dropping a course presented primarily in each format, f.) predicted likelihood of watching an entire lecture across different formats. The audio+instructor condition differs significantly across all subjective reports from audio alone and text alone, but is not rated differently from audio+text. g.) The effect of lecture format on participants' Judgments of learning (JOL). The main effect of JOL is driven by a difference between audio only and text only formats. Error bars reflect standard error of the mean (SEM).). * indicates $p < 0.05$, ** indicates $p < 0.001$.

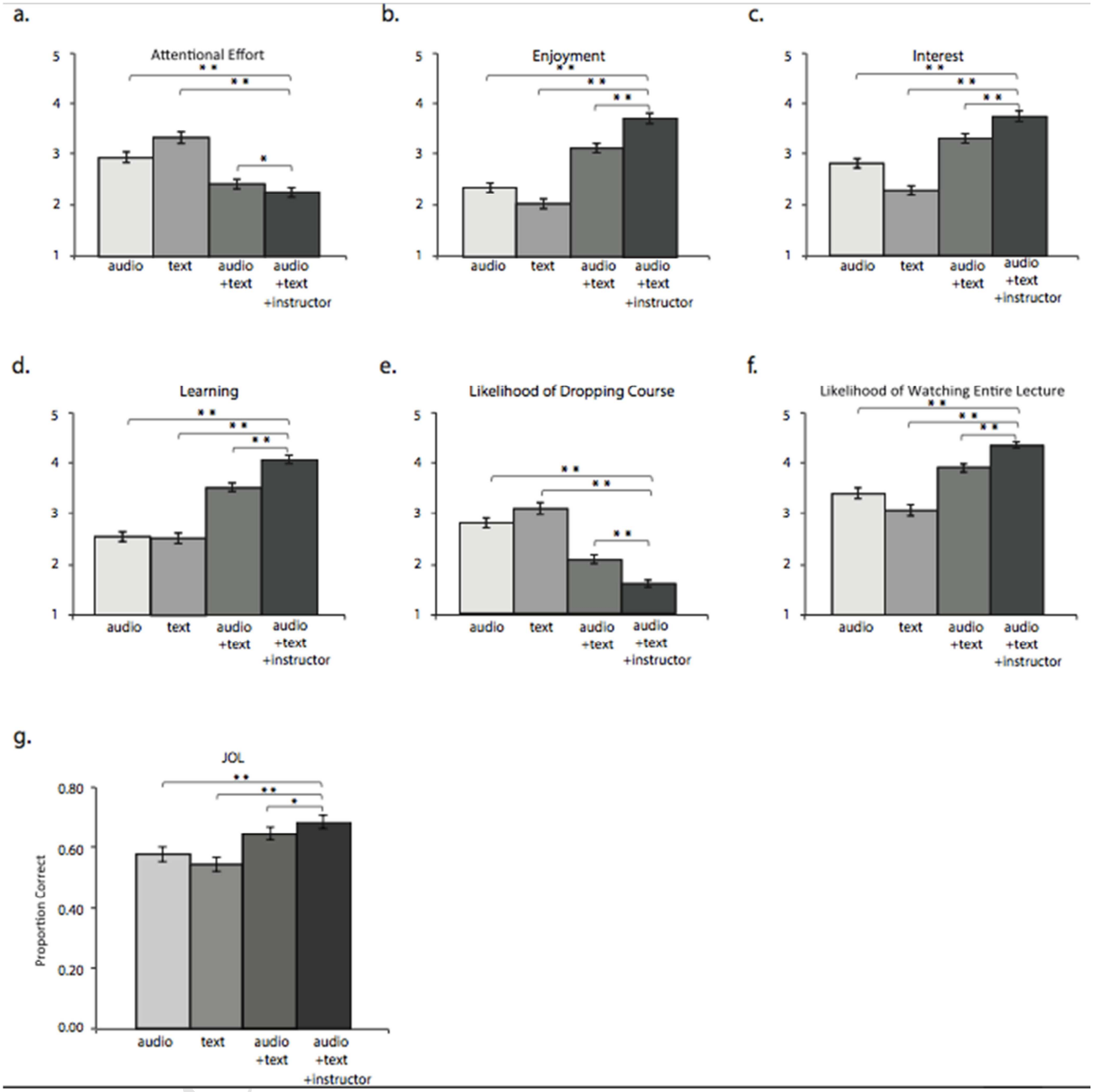

Figure 7. Experiment 4: Subjective ratings (Likert Scale, 1-5, ranging from a lot/very to very little). a.) amount of effort maintaining attention, b.) degree of enjoyment watching lecture across different formats, c.) level of interest watching lecture across different formats, d.) how much learning is facilitated by each format, e.) predicted likelihood of dropping a course presented primarily in each format, f.) predicted likelihood of watching an entire lecture across different formats. The audio+text+instructor condition, is rated significantly higher across all subjective reports relative to the other formats. g.) Participants' Judgments of learning (JOL) are also significantly higher for the audio+text+instructor format relative to the other 3 formats. Error bars reflect standard error of the mean (SEM). * indicates $p < 0.05 > 0.001$; ** indicates $p < 0.001$.

Highlights

- It is unclear if visuals of instructors (video) in online lectures aid learning
- In 4 experiments videos with instructors were contrasted with other lecture formats
- Comprehension, attention (mind wandering) and subjective evaluations were measured
- The instructor's presence impairs comprehension, yet learners far prefer this format
- This research highlights a tension between learning and liking in online education