Instructor enthusiasm in online lectures: how vocal enthusiasm impacts student engagement, learning, and memory

Jeremy Marty-Dugas¹,²*, Maya Rajasingham², Robert J. McHardy³, Joe Kim² and Daniel Smilek¹

¹Vision & Attention Lab, Department of Psychology, University of Waterloo, Waterloo, ON, Canada, ²EdCog Lab, Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada, ³Hearts & Minds Lab, Department of Psychology, University of Manitoba, Winnipeg, MB, Canada

Across two studies we investigated the impact of instructor enthusiasm on student attentional engagement during an online undergraduate lecture, as well as their memory for lecture content and their motivation to watch additional lecture videos on the same topic (Study 2 only). In both studies participants were randomly assigned to watch a 22-min lecture, delivered with either high or low vocal enthusiasm by the instructor. Subjective ratings of instructor enthusiasm/energy confirmed that in both studies the manipulation of instructor enthusiasm was effective. More importantly, in both studies we found that students in the high enthusiasm condition were consistently more engaged over the course of the lecture compared to those in the low enthusiasm condition, and that overall, reports of engagement increased together with ratings of instructor enthusiasm. However, we found no evidence that instructor enthusiasm influenced quiz performance in either study. Nevertheless, Study 2 showed that those in the high enthusiasm condition were more motivated to watch the next lecture than those in the low enthusiasm condition. These findings make an important contribution to the study of online learning and indicate that instructor enthusiasm may be a viable strategy to increase student engagement and motivation in online courses.

KEYWORDS
engagement, enthusiasm, attention, online learning, education, memory

Introduction

Increasingly, learning is taking place online (Koskal, 2020a,b), a trend that has been accelerated by the Covid-19 pandemic (Li and Lalani, 2020). While online learning is touted as having numerous benefits, mostly derived from improving accessibility, as online lectures and course materials can be accessed at any time from any place, this form of learning also has some notable limitations. One problem is a high rate of inattentiveness during online lectures, which tends to increase as a lecture progresses (Risko et al., 2012, 2013; Kane et al., 2017; Wammes and Smilek, 2017). Inattention during online lectures is particularly problematic because it has been associated with impaired comprehension and learning (Risko et al., 2012, 2013; Szpunar et al., 2013a,b; Kane et al., 2017). Given the central role of attention in learning,
it is important to consider the various factors that may influence attentional engagement during online learning.

Research on learning both online and in person has revealed several factors that influence attentional engagement. Inattention in learning contexts has been shown to be reduced by, interpolated testing (Szpunar et al., 2013b; Schacter and Szpunar, 2015), pretesting (Pan et al., 2020), and the inclusion of short breaks that involve physical activity (Mahar, 2011; Fenesi et al., 2018). Implementing mindfulness exercises (Wilson and Dixon, 2010) and banning the use of distracting technology in class also seem to improve student attentional engagement (Aguilar-Roca et al., 2012; Glass and Kang, 2019). In the present investigation, we explore another potential way to increase student attention (i.e., engagement) during online undergraduate lectures—namely, increasing the level of enthusiasm of the teacher or instructor.

Instructor enthusiasm and student experiences

The notion that an enthusiastic teacher “can make all the difference” is a popular idea among both the general population and educators (e.g., Barr, 1929; Hamachek, 1975, p. 303; Minor et al., 2002), so much so that teachers will even feign enthusiasm (Burić, 2019), despite the emotional cost of doing so (Taxer and Frenzel, 2018). Teacher enthusiasm can be conveyed in multiple ways, including via energy levels, vocal dynamics, eye gaze, facial expressions, gestures, movements, and word selection (see Rosenshine, 1970; Collins, 1976, 1978). The way enthusiasm can be expressed likely depends on whether the class is in-person or online. Instructors who teach in person can communicate their enthusiasm through their use of the physical space, such as walking down the aisles or towards students who ask questions, and may be able to modulate their enthusiasm levels on the fly in response to visual feedback from their students (e.g., looks of interest, confusion, or sleepiness). In contrast, those teaching online (whether live or using a pre-recorded video lecture) have to rely more heavily on vocal dynamics and word choices, especially if relying on their voice alone (e.g., power point slides with voiceover).

Setting aside learning and attentional engagement for a moment, research has found, consistent with popular belief, a positive association between instructor enthusiasm and a variety of student experiences (Keller et al., 2016). For example, higher instructor enthusiasm tends to be related to higher academic self-efficacy (Zhang, 2014), higher levels of intrinsic motivation (Valentín et al., 2022), greater interest in the course content (Patrick et al., 2000; Keller et al., 2014; Kim and Schallert, 2014), lower class-related boredom (Cui et al., 2017a,b; Cui et al., 2022), and even greater academic honesty (Oroz et al., 2015). Instructor enthusiasm also seems to confer a more positive mental and affective state (Frenzel et al., 2019), such as increased feelings of “vitality” (i.e., energy, alertness and wakefulness).

A good example of a study examining instructor enthusiasm on student experiences in an online context is one reported by Frenzel et al. (2019), which examined university students watching a short video lecture (i.e., six minutes). The short instructional video featured an instructor (actually a professional actor) standing at a podium, or using visual aids, to explain the effectiveness of various learning techniques. Critically, instructor enthusiasm was varied only by the way the instructor delivered the material. In the more enthusiastic condition, the actor aimed to display a high degree of energy, gesticulate, be verbally vigorous, facially expressive and to maintain eye contact with the camera. In contrast, in the less enthusiastic condition, the actor aimed to do the opposite (i.e., display low energy etc.; see Frenzel et al., 2019, p. 258). The presentations script was the same across conditions. The results showed that students reported more enjoyment and less boredom while watching the video in the more enthusiastic condition.

Instructor enthusiasm and learning

Several studies have examined the impact of instructor enthusiasm on students’ memory and comprehension (Mastin, 1963; Coats and Smidchens, 1966; Marsh, 1984; Moë et al., 2021), reporting somewhat mixed results (Keller et al., 2016). Starting with the positives, there are several studies showing that increased instructor enthusiasm leads to improved learning. For instance, in a study examining the impact of enthusiasm in junior high classrooms, Mastin (1963) found that student performance was significantly higher (i.e., greater memory for content as assessed by a multiple-choice test) for lessons in which the teacher delivered the material enthusiastically, compared to an “indifferent” manner. In another study (Coats and Smidchens, 1966), university students who listened to a “dynamic” speaker had significantly better recall of the 10-min speech (as assessed by a multiple-choice test) compared to those who listened to a “static” speaker. Along similar lines, Stewart (1989) experimentally manipulated instructor enthusiasm (enthusiastic vs. non-enthusiastic content delivery) in live lectures which were delivered as part of a real university course. When students were permitted to take and review their notes prior to the test, those in the enthusiastic condition performed significantly better—although this pattern reversed when students were not allowed to review their notes1. More recently, it was found that participants who had an enthusiastic instructor while completing an interactive programming lesson online performed better on a subsequent test (Liew et al., 2017, 2020); in this investigation, the instructor2 in the enthusiastic condition also explicitly stated they were enthusiastic about the topic. Finally, Moë (2016) found that children assigned to listen to an enthusiastic reader had greater recall of the material than children assigned to listen to a non-enthusiastic reader (see also Moë et al., 2021).

Similar positive influences of instructor enthusiasm on learning have been reported in the context of video-based learning (Marsh, 1984; Bege et al., 2020; Lawson et al., 2021 Exp 2; Valentín et al., 2022; Bege and Schneider, 2023). For example, a re-analysis conducted by Marsh (1984) of a series of studies that made use of video lectures in the 1970s (Naftulin et al., 1973; Ware and Williams, 1975; Williams and Ware, 1976) showed that university students who viewed a video with a more enthusiastic presenter performed better on a memory task.

---

1 There was no observable effect when students were instructed to “listen only.”

2 In this case, the instructor was a digital avatar paired with a human voice narration.
recognition test delivered immediately following the lecture. As a more recent example, Valentín et al. (2022) showed elementary school children videos of stories being presented by teachers who displayed either high or low enthusiasm. Learning was assessed by indexing students’ abilities to describe the ideas in the stories, retrieving details, make inferences and interpret the contents. Results showed learning was overall higher for students viewing videos of more enthusiastic teachers than the more neutral teachers.

However, in contrast to the foregoing, several studies examining in-person learning have indicated that instructor enthusiasm has either no effect, or even has a detrimental effect on learning. For example, McKinney et al. (1983) trained six teachers to vary their enthusiasm to be low, medium, or high. They found no effect of teacher enthusiasm on the achievement of fourth-grade students, and further, received informal reports of a detrimental effect of high enthusiasm on classroom management in the high enthusiasm condition. Larkins and McKinney (1982) reported a similar effect in seventh-grade students (Study 1), although another experiment (Study 2) found that both the high and moderate enthusiasm conditions outperformed the low enthusiasm condition. In a similar design with first-grade students, McKinney et al. (1984a) reported a negative effect of high enthusiasm, while another study also found no effect of enthusiasm on achievement in three and four-year-old children (Burts et al., 1985). Moving to older age groups, McKinney et al. (1984b) did not find any differences in achievement among undergraduate students assigned to low, medium and high instructor enthusiasm conditions, while another study found that enthusiastic video instructions did not impact performance on a subsequent memory task, compared to neutral instructions’ (Motz et al., 2017).

Similar null effects of enthusiasm on learning have been found in the context of university-level video lectures. For example, in several studies that involved participants watching a 10-min video lecture on statistics (Horovitz and Mayer, 2021; Lawson et al., 2021 Exp 1; Lawson and Mayer, 2022b), the instructor displayed one of several emotions; happiness, contentedness, boredom, or frustration. While participants were able to reliably identify the emotion displayed by the instructor, there was no significant effect on test performance on either immediate (Horovitz and Mayer, 2021; Lawson et al., 2021, Exp 1) or delayed (Lawson and Mayer, 2022b) posttests.

**Instructor enthusiasm and attention**

Perhaps most important for present purposes are several studies that have suggested instructor enthusiasm can positively affect engagement, as indicated by student attention to learning material (Coats and Smidchens, 1966; Bettencourt et al., 1983; Burts et al., 1985; Moè et al., 2021), and that increased attention to the instructor may be one mechanism by which enthusiasm leads to improved recall. For example, Moè et al. (2021), investigated this in a series of experiments in fourth and fifth graders by varying the enthusiasm displayed by the instructor when reading a three-minute passage out loud. They found that compared to students in the non-enthusiastic condition, those in the enthusiasm condition were more attentive, as assessed by both self-reports and behavioral coding (i.e., time spent looking at the reader), and had greater recall of the stories. Interestingly, they also noted that the effect of high vs. low enthusiasm was negated when students were asked to perform a concurrent visual search task (Experiments 2 and 3), suggesting that attention is a necessary component for an effect of enthusiasm on recall to occur.

Also worth noting are several studies that examined the effects of an instructor's displayed emotions in the context of video lectures at a university level. Although they did not investigate enthusiasm per se, some of the conditions could be construed as varying in enthusiasm (i.e., happiness vs. boredom). Also, while these studies did not seek to assess attention directly, they measured participants’ experiences that are attention adjacent. For example, participants were asked to report on their motivation to pay attention, and whether they would like to receive similar lessons in the future, which could be construed as a decision about what to attend to in the future (Horovitz and Mayer, 2021; Lawson et al., 2021; Lawson and Mayer, 2022b). In general, it was found that lectures from instructors displaying a positive emotion led to increased scores on these measures (Horovitz and Mayer, 2021; Lawson et al., 2021). Furthermore, participants also tended to rate instructors as more engaging (i.e., ratings of the degree to which the instructor was friendly, expressive, enthusiastic, motivating and entertaining) when they displayed active and positive emotions (Lawson et al., 2021; Lawson and Mayer, 2022a).

Importantly, however, the questionnaires in these studies were always delivered after participants completed the posttest, rather than immediately after the video lecture. As the posttest was generally more than twice the length of the video lecture’ (Lawson et al., 2021), and sometimes delivered after a week's delay (e.g., Lawson and Mayer, 2022b), there is potential for participant's assessments of their experience during lecture to have been influenced by their experience with the test. Further, given that longer delays after an event tend to lead to greater forgetting (Ebbinghaus, 1964; Schacter, 1999; Roediger and Karpicke, 2006), there is a clear need for studies that specifically assess the effect of enthusiasm during online lectures on attentional engagement using more immediate assessments.

When considering the literature, many studies of the effects of enthusiasm on attention have focused on shorter time periods, younger children, or used only indirect measures of attention, which limits the conclusions that can be drawn about the impact of teacher enthusiasm on students’ attention. For example, in some studies, the claim that instructor enthusiasm influences student attention is based on informal reports from the classroom teacher (Burts et al., 1985), or it is inferred that the students must have had greater attention to the lecture based on their subsequent memory performance (Coats and Smidchens, 1966). In other cases, the studies employ measures that do not clearly assess attentional engagement (Bettencourt et al., 1983; Huangfu et al., 2022, 2024), or assess attention in an indirect or peripheral way (Liew et al., 2017, 2020; Beege et al., 2020; Horovitz and Mayer, 2021; Lawson et al., 2021; Lawson and Mayer, 2022a; 2024).

---

3 However, participants who saw enthusiastic instructions did attempt to complete more word pairs (i.e., they attempted to remember more items during the task), which could be interpreted as a positive effect of enthusiasm on motivation.

4 The posttest was self-paced and took an average of 26min, while the video lecture was 10min long.
because it was not directly related to their research goals. Other studies have focused on attention in young children over short time periods (Moë, 2016; Moë et al., 2021). In sum, relatively few studies have examined how enthusiasm impacts attention in online/video lectures, and many of the aforementioned studies (particularly those that use video lectures) are brief, lasting only ten minutes or less. As such, one must exercise caution when attempting to draw conclusions about how, for example, enthusiasm impacts attention over longer periods of time, or the effectiveness of enthusiasm on the attention of college and university students viewing video lectures.

The present study

Our overarching interest was to examine how lecturer enthusiasm influences students’ attentional engagement with online lecture content as the lecture unfolds over time. The present work builds on the extant literature in several ways: First, and most important, given that few studies have examined how enthusiasm impacts attention (i.e., engagement defined as attention to the lecture), we focused on assessing participants’ experience of attention using assessments of absorption, immersion, and deep effortless concentration (which can be characterized as peak attention; Tellegen and Atkinson, 1974; Csikszentmihalyi, 1978; Reed et al., 2002; Csikszentmihalyi and Nakamura, 2010; Peifer et al., 2014; Marty-Dugas and Smilek, 2019).

Second, given that attentional engagement during video lectures tends to decline over time on task (Risko et al., 2012; Farley et al., 2013; Risko et al., 2013; Kane et al., 2017), we employed a video lecture that is substantially longer than most prior studies examining enthusiasm in this format. Critically, this allowed us to assess whether the effect of enthusiasm on attention, if any, varies in degree over the course of the lecture. Third, we used a lecture drawn from an actual university course and delivered by a regular instructor of the course (rather than a professional actor), thus maximizing ecological validity. And finally, we employed a lecture with a “voiceover and slides” style, which is a common format for online learning (Chen and Thomas, 2020).

In two studies (i.e., two samples), participants viewed a 22-min lecture on the topic of sleep and circadian rhythms, presented in an online format, which was followed by a 16-item quiz on the lecture content. The video included an auditory narration by the instructor and visual PowerPoint slides with some animations; the instructor was not visibly present in the video. Thus, we focused specifically on the effect of enthusiasm as conveyed by the instructor’s voice. We created enthusiastic and non-enthusiastic conditions by altering the instructor’s level of vocal enthusiasm (i.e., the amount of variability in pitch and intonation; see video clips on OSF). Visually, the experience in both conditions was held constant. By using the same lecture script, we also ensured that the audio content—in terms of what was said—was virtually identical across conditions.

In both studies, we measured participant’s assessment of the instructor’s enthusiasm, and their own level of engagement, as well as participants’ learning. Participants’ judgments of instructor enthusiasm and their own level of engagement were measured with experience sampling probes (i.e., thought probes) that were presented intermittently throughout the lecture. When measuring engagement we assessed the frequency with which student experience deep effortless concentration, as well as feelings of immersion, and absorption. In addition, in both studies we indexed learning via a multiple-choice test on the lecture content which was administered after the lecture was presented. Finally, in Study 2, we investigated the impact of enthusiasm on student’s motivation to access the next lecture in the series.

This study design allowed us to address several research aims. First, our primary aim was to examine whether variation in instructor enthusiasm leads to a difference in students’ level engagement (i.e., their experience of attention), and whether this effect changes over the various probes presented at different time points during the lecture. On the one hand, the effect of enthusiasm may only emerge later in the lecture, when participant’s attention is more likely to lapse. On the other hand, the effect may occur primarily at the beginning of the lecture, when participants are more attentive and better able to detect instructor enthusiasm. Second, we sought to explore the relation between participants’ ratings of instructor enthusiasm and their self-reported levels of engagement. Based on prior findings, it was expected that as ratings of instructor enthusiasm increase, so does participant engagement with the lecture. Third, we explored the impact of the enthusiasm manipulation on students’ quiz performance as a means of assessing whether enthusiasm had any impact on students’ memory for lecture content, and anticipated that memory performance should be greater in the enthusiastic condition relative to the non-enthusiastic condition. Finally, in our second study we tested whether instructor enthusiasm influences participants’ motivation to re-engage in subsequent similar course content. The rationale was that increasing instructor enthusiasm might increase the likelihood that students would want to watch course content in the future.

Transparency and openness

We note that in the present investigation, we present two studies, the first of which (Study/Sample 1) served as an exploratory sample and the second of which (Study/Sample 2) served as a confirmatory sample. The procedure for both samples was identical, except for an additional question about participant motivation included in Study 2, which was presented at the end of the lecture, but prior to the final quiz, so that answers to the motivation question were not affected by quiz performance. After completion of the exploratory data analysis (i.e., Study 1), we posted our R analysis code, as well as a draft of the manuscript and results, to the associated OSF page. This served as

---

5 Further, prior work (Lawson and Mayer, 2022a) found that participants who viewed a lecture which had a visual of the instructor, did not perform significantly better at identifying instructor emotions compared to a group who watched a lecture that used the instructor’s voice only.

6 Our prediction for the motivation question, as well as the planned analyses to assess this question were also pre-registered.

7 Specifically, the draft of the current manuscript from the intro through the data analysis plan for Study 2. Alterations have been made according to reviewer recommendations and journal requirements.

8 https://osf.io/t9drh/
the pre-registration of our analyses for Study 2, where we predicted we would replicate our findings, and conducted confirmatory analyses by repeating these same analyses, using the same R code.

Study 1

Methods

Participants

Participants were recruited from Psychology courses at the University of Waterloo in exchange for partial course credit. Using a website for participant pool management (i.e., SONA), students elected to participate in the present study by selecting from a list of possible studies and signing up for online timeslots. After providing informed consent, participants were asked to answer a brief demographic questionnaire to collect data on aspects such as age, gender and racial identity. Participants were also asked to indicate their interest in and familiarity with the topic of circadian rhythms. Participants who completed the study more than once, did not respond to all the probes, and/or scored equal to or less than chance performance (25%) on the end-of-lecture quiz were removed from the sample. After the data screening procedure, 139 participants (104 Female, 29 Male, 4 Non-binary, 1 Genderqueer, 1 prefer not to answer)—69 in our high enthusiasm condition, and 70 in our low enthusiasm condition—were included in the final sample.

Measures

Professor enthusiasm

To quantify professor enthusiasm, participants were asked to rate how enthusiastic/energetic they found the professor on a scale of 0 (not at all enthusiastic/energetic) to 100 (extremely enthusiastic/energetic) to provide a wide range of possible answer options.

Engagement

Engagement (i.e., attention to the lecture) was quantified using five self-report items which featured questions such as “I was totally absorbed by the lecture” and “I was able to completely focus without straining to pay attention.” Prior work has used analogues of items one and two to measure the concepts of immersion and absorption (Engeser and Rheinberg, 2008; Peifer et al., 2014; Thissen et al., 2018, 2021), and items three to five to assess deep effortless concentration (i.e., flow; Marty-Dugas and Smilek, 2019). Each item was rated on a scale ranging from 0 (almost never) to 100 (almost always). For an example of the probe, see Appendix 1. The probe was presented eight times to each participant, precise timings for each probe are listed on OSF.9 An engagement score was calculated by finding the average of the five items at each probe for each participant.

Multitasking

To quantify multitasking participants were asked to report how frequently they used a digital device during the lecture on a scale ranging from 0 (almost never) to 100 (almost always). Additionally, we collected data on “blur times” which indexed how much time participants had a window other than the experimental window as an active tab on their computer. These data are reported elsewhere.

Memory performance

A 16-item multiple (MC) choice quiz, testing concepts from various points throughout the lecture, was implemented to quantify memory performance. For each participant, roughly half of these questions were based on content that was closely followed by a thought probe, while the other half of the items were not. For those participants in Version A, the items closely followed by a probe were (1, 2, 5, 7, 9, 11, 13, 15), while for those in Version B, items (3, 4, 6, 8, 10, 12, 14, 16) were closely followed by the probe. Each question had four options which participants were instructed to choose from, except for item 1 which erroneously included a fifth answer option. For the full list of questions, see Appendix 2.

Procedure

Participants were randomly assigned to one of two conditions—high enthusiasm or low enthusiasm. For each condition, participants viewed an asynchronous 22-min lecture on the topic of circadian rhythms and sleep. The lectures presented in both conditions were virtually identical in terms of their content, as they made use of the exact same slides and followed the same script. The lectures differed in terms of the professor’s delivery/tone of voice (i.e., either high or low enthusiasm). Brief audio clips of each condition, as well as the slides and script, are available on the OSF (see text footnote 8).

To assess the effect of professor enthusiasm on student engagement, eight thought probes (i.e., experience sampling probes) were presented intermittently throughout the lecture. At each probe screen, participants were instructed to rate the frequency with which they experienced each of the listed experiences since the last time they saw the probe screen10. In order to sample participants’ experiences from a greater percentage of the lecture, we used two sets of probe timing (e.g., Version A and Version B; see Figure 1) to which participants were randomly assigned. Thus, across Version A and B, participant experiences were collected from a total of 16 different time points. The information presented at these timepoints was used to create the 16-item multiple choice test of memory for lecture content. All participants were presented with the same 16 items on the test. By using two sets of probe timings and collapsing them together, it allowed us to sample participants’ experiences at more timepoints throughout the lecture, without interrupting any individual too often. Doing so also helped to assuage concerns that any effects of enthusiasm were unique to the specific time points the participants were probed at11. Following the conclusion of the lecture, all

9 https://osf.io/tgp9n

10 On the first probe participants were instructed to rate their experiences “since the beginning of the experiment.”

11 Precise timings for the information the questions were based on can be found at: https://osf.io/tgp9n. The probes were presented at the end of these time windows.
participants were presented with the same untimed 16-item multiple choice quiz on the lecture content.

Data cleaning

The full data cleaning procedure can be found on the associated OSF page in our R Code or in the html document. The full raw data, including those participants who were removed prior to analyses, can be found there as well. As a first step, we removed pilot data all of whom accessed the data from outside the online recruitment system12. To get our measures of engagement, for each participant we averaged the scores from the five engagement items at each probe. This left each participant with eight scores, one for each probe.

Next, we checked to ensure there were no duplicate participants in the data. We identified six duplicate IDs indicating that these participants had participated twice. As we had no way to ensure which data was associated with the participants first time completing the experiment, these six participants were removed from both the aggregate and the probe data.

Following this we examined our variables of interest using violin plots. One participant in the non-enthusiastic condition was identified as an extreme outlier using the identify_outliers() function from the rstatix package (Kassambara, 2023a,b) and based on visual inspection (see R code on OSF for the visual). 13 participants who performed quiz (i.e., 25% or less) were removed due to concerns about random responding adding noise to the data. Nine participants who had self-report data, but no quiz data were removed.

Results and discussion

Statistical analyses were conducted using R 4.3.0 (R Core Team, 2023). The ANOVAs were conducted using the ez package (Lawrence, 2016) and follow-up t-tests were conducted using the rstatix package (Kassambara, 2023a,b). Further, bootstrapped 95% CIs of effect size using the bias-corrected and accelerated method (Efron, 1987; Kirby and Gerlanc, 2013) were generated using rstatix13 (Kassambara, 2023a,b). Data cleaning and visualizations made use of the tidyverse (Wickham et al., 2019), ggpubr (Kassambara, 2023a), data summary (Arel-Bundock, 2022) and gt (lannone et al, 2023) packages. The raw data and the full code used to clean and analyze the data for is available on the study OSF page: https://osf.io/t9drh/.

Manipulation check

We began by analyzing ratings of instructor enthusiasm to determine whether or not our manipulation of instructor enthusiasm had the desired effect. To do so, we conducted a 2×2×8 mixed ANOVA with Condition (High or Low Enthusiasm) and Version14 (A or B) as the between-subjects factors and Time (Probes 1 through 8) as the within-subjects factor. Mauchley’s test indicated that the assumption of sphericity was violated W = 0.10, p < 0.05, so a Greenhouse Geiser correction was applied. As can be seen in Table 1, the omnibus ANOVA revealed a significant main effect of Condition F(1, 135) = 115.00, p < 0.001, η²p = 0.41, such that there was a large effect of condition on participants ratings of the instructors enthusiasm/energy. There was no significant main effect of Version F(1, 135) = 0.00, p = 0.964, η²p = 0.00, nor a significant main effect of Time F(3.67, 495.19) = 1.04, p = 0.384, η²p = 0.00. Further, there were no significant interactions between any of the variables in the model (see Table 1). As there was no effect of Version, we collapsed the means across version in Table 1 to more clearly highlight our comparison of interest. A full table of the non-collapsed means can be found in Supplementary Tables on the associated OSF page.

While the ANOVA indicated there was no significant effect of Time, we nonetheless computed bootstrapped 95% CIs of Cohen’s d at each timepoint to better illustrate the effect of the manipulation. As can be seen in Table 2, the effect of condition was consistently large across the duration of the lecture (see also New Figure - Ratings of Enthusiasm). Taken together, these results indicate that our manipulation of instructor enthusiasm was successful and consistently present throughout the lecture.

The effect of instructor enthusiasm on participants’ attentional engagement

Next, we examined the effect of enthusiasm on participants’ ratings of their attentional engagement during the lecture, which was our main question of interest. We conducted a 2×2×8 mixed ANOVA on engagement with Condition (High or Low Enthusiasm) and Version (A vs. B) as the between-subjects factors and Time (Probes 1 through 8) as the within-subjects factor.

Mauchley’s test indicated that the assumption of sphericity was violated W = 0.13, p < 0.05, so a Greenhouse Geiser correction was applied. As can be seen in Table 3, the omnibus ANOVA revealed a significant main effect of Condition F(1, 135) = 28.88, p < 0.001, η²p = 0.14, such that there was a large effect of condition on participant engagement. There was no significant main effect of Version F(1, 135) = 0.17, p = 0.683, η²p = 0.00. There was a significant main effect of

---

12 Initially, our plan was to collect additional data from a non-student population using the lab mailing list. Only one person participated via this method. Further, due to a coding error all participants from outside our online recruitment system (including the pilot testers) were assigned the same random ID, so this participant was removed along with the pilot data.

13 The default setting is to use the percentile method. We changed this by setting ci.type = “bca”.

14 Recall that version indicates which set of probe timings participants received.
Time $F(3.95, 533) = 2.95, p < 0.05, \eta^2_g = 0.00$. Further, there was a significant interaction between Condition and Time $F(3.95, 533) = 2.74, p < 0.05, \eta^2_g = 0.00$. To investigate the interaction, we conducted post-hoc Welch’s t-tests with a Bonferroni correction for multiple comparison using the rstatix package (Kassambara, 2023a,b). Further, we also used this package to calculate Cohen’s $d$ to get a measure of effect size, as well as generated bootstrapped 95% CIs ($n=10,000$) of the effect size. These results are presented in Table 4. Note, as there was no effect of Version, we collapsed the means across Version in Table 4 to more clearly highlight our comparison of interest (A full table of the non-collapsed means can be found in Supplementary Tables on OSF).

As can be seen in the table, the effect of condition was significant at each probe, such that those in the high enthusiasm condition were significantly more engaged than those in the low enthusiasm condition. Further, as can be seen from examining the effect size estimates, the effect of enthusiasm on engagement generally ranged from a moderate to large effect throughout the experiment—interestingly, the estimate of the effect was smallest at Time 1 and increased at Time 2, after which it remained relatively stable. This would seem to indicate that the effect of enthusiasm on engagement was least pronounced in the earliest part of the lecture. Regardless of any changes in the size of the effect, these results indicate that enthusiasm has an impact on student engagement throughout the lecture, such that those in the more enthusiastic lecture (i.e., high enthusiasm) were consistently more engaged.

### The relation between ratings of instructor enthusiasm and attentional engagement

Next, we examined whether participants’ perceptions of instructor enthusiasm were related to levels of attentional engagement. Because condition had a significant effect on both enthusiasm and engagement ratings, we computed a separate correlation between enthusiasm and engagement ratings in each instructor enthusiasm condition (i.e., high and low enthusiasm). As can be seen in Figure 2, there was a moderate, positive correlation between ratings of enthusiasm and participant engagement in both conditions, such that those who found the professor more enthusiastic/energetic tended to indicate that they were more engaged during the lecture. Interestingly, the correlations in the two conditions were similar in magnitude. However, the positive relation between ratings of enthusiasm and engagement appears to tail off at the top end of enthusiasm ratings, perhaps suggesting that for some students, there is a point at which increased professor enthusiasm no longer leads to greater engagement, and this occurs even before maximal levels of engagement are reached.

### The effect of instructor enthusiasm on quiz performance

To determine whether the effect of enthusiasm was sufficient to impact participants’ memory for the lecture content (i.e., assessed via quiz performance), we also conducted a 2x2 ANOVA with Condition (High vs. Low Enthusiasm) and Version (A vs. B) as between-subjects factors. There was no significant main effect of Condition $F(1, 135) = 2.67, p = 0.104, \eta^2_g = 0.02$, Version $F(1, 135) = 0.03, p = 0.862, \eta^2_g = 0.00$, or the interaction $F(1, 135) = 1.48, p = 0.226, \eta^2_g = 0.01$. The full table of means can be found in Supplementary Tables on OSF. Once again, we examined the effect size of the difference between the two conditions (collapsing across Version), and generated bootstrapped confidence intervals.

### Table 1: Sample 1 ANOVA results for ratings of instructor enthusiasm

<table>
<thead>
<tr>
<th>Predictor</th>
<th>df (num)</th>
<th>df (den)</th>
<th>Epsilon</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_g$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>1.00</td>
<td>135.00</td>
<td>115.00</td>
<td>0.000</td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>1.00</td>
<td>135.00</td>
<td>0.00</td>
<td>0.964</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Condition x Version</td>
<td>1.00</td>
<td>135.00</td>
<td>0.65</td>
<td>0.421</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Time (probe)</td>
<td>3.67</td>
<td>495.19</td>
<td>0.52</td>
<td>1.04</td>
<td>0.384</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition x Time (probe)</td>
<td>3.67</td>
<td>495.19</td>
<td>0.52</td>
<td>0.51</td>
<td>0.713</td>
<td>0.00</td>
</tr>
<tr>
<td>Version x Time (probe)</td>
<td>3.67</td>
<td>495.19</td>
<td>0.52</td>
<td>1.58</td>
<td>0.184</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition x Version x Time (probe)</td>
<td>3.67</td>
<td>495.19</td>
<td>0.52</td>
<td>1.77</td>
<td>0.139</td>
<td>0.00</td>
</tr>
</tbody>
</table>

$df_{num}$ indicates degrees of freedom numerator; $df_{den}$ indicates degrees of freedom denominator. Epsilon indicates Greenhouse–Geisser multiplier for degrees of freedom, $p$-values and degrees of freedom in the table incorporate this correction. $\eta^2_g$ indicates generalized eta-squared. Table generated using APATables (Stabley, 2021).

### Table 2: Sample 1 descriptive statistics, post-hoc tests and 95% CIs for ratings of instructor enthusiasm

<table>
<thead>
<tr>
<th>Enthusiasm</th>
<th>Mean (SD)</th>
<th>Mean (SD)</th>
<th>$t^a$</th>
<th>df$^a$</th>
<th>$p$ adj$^b$</th>
<th>d</th>
<th>95% CI$^c$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>20.16 (17.31)</td>
<td>59.32 (26.31)</td>
<td>−10.35</td>
<td>117.34</td>
<td>0.00</td>
<td>1.76</td>
<td>[1.29, 2.23]</td>
</tr>
<tr>
<td>High</td>
<td>22.40 (22.32)</td>
<td>60.17 (21.89)</td>
<td>−10.18</td>
<td>137.00</td>
<td>0.00</td>
<td>1.73</td>
<td>[1.23, 2.24]</td>
</tr>
<tr>
<td></td>
<td>22.21 (23.10)</td>
<td>59.96 (22.83)</td>
<td>−9.69</td>
<td>137.00</td>
<td>0.00</td>
<td>1.64</td>
<td>[1.18, 2.11]</td>
</tr>
<tr>
<td></td>
<td>21.93 (23.27)</td>
<td>59.94 (24.26)</td>
<td>−9.43</td>
<td>136.57</td>
<td>0.00</td>
<td>1.60</td>
<td>[1.12, 2.07]</td>
</tr>
<tr>
<td></td>
<td>22.06 (23.03)</td>
<td>60.84 (25.33)</td>
<td>−9.44</td>
<td>135.39</td>
<td>0.00</td>
<td>1.60</td>
<td>[1.14, 2.08]</td>
</tr>
<tr>
<td></td>
<td>21.40 (22.16)</td>
<td>59.87 (26.49)</td>
<td>−9.28</td>
<td>132.20</td>
<td>0.00</td>
<td>1.58</td>
<td>[1.11, 2.07]</td>
</tr>
<tr>
<td></td>
<td>22.27 (23.91)</td>
<td>64.06 (26.45)</td>
<td>−9.77</td>
<td>135.21</td>
<td>0.00</td>
<td>1.66</td>
<td>[1.18, 2.15]</td>
</tr>
<tr>
<td></td>
<td>20.94 (23.90)</td>
<td>60.71 (25.93)</td>
<td>−9.40</td>
<td>135.76</td>
<td>0.00</td>
<td>1.59</td>
<td>[1.12, 2.11]</td>
</tr>
</tbody>
</table>

Welch’s t-tests were conducted.

*a Bonferroni correction for multiple comparisons was applied.

*b 95% CIs of effect size generated via bootstrapping ($n=10,000$) with the bca method.
TABLE 4 Sample 1 descriptive statistics, post-hoc tests and 95% CIs for ratings of engagement.

<table>
<thead>
<tr>
<th>Time</th>
<th>Low enthusiasm</th>
<th>High enthusiasm</th>
<th>t</th>
<th>df</th>
<th>p.adj</th>
<th>d</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>34.67</td>
<td>46.61</td>
<td>−2.92</td>
<td>136.97</td>
<td>0.03</td>
<td>0.50</td>
<td>[0.14, 0.84]</td>
</tr>
<tr>
<td>2</td>
<td>30.88</td>
<td>54.08</td>
<td>−5.67</td>
<td>137.00</td>
<td>0.00</td>
<td>0.96</td>
<td>[0.56, 1.38]</td>
</tr>
<tr>
<td>3</td>
<td>29.18</td>
<td>51.23</td>
<td>−5.08</td>
<td>136.94</td>
<td>0.00</td>
<td>0.96</td>
<td>[0.52, 1.31]</td>
</tr>
<tr>
<td>4</td>
<td>29.24</td>
<td>51.00</td>
<td>−4.95</td>
<td>136.83</td>
<td>0.00</td>
<td>0.84</td>
<td>[0.46, 1.22]</td>
</tr>
<tr>
<td>5</td>
<td>27.07</td>
<td>49.88</td>
<td>−5.37</td>
<td>136.77</td>
<td>0.00</td>
<td>0.91</td>
<td>[0.46, 1.21]</td>
</tr>
<tr>
<td>6</td>
<td>26.59</td>
<td>48.16</td>
<td>−4.90</td>
<td>136.91</td>
<td>0.00</td>
<td>0.83</td>
<td>[0.40, 1.15]</td>
</tr>
<tr>
<td>7</td>
<td>27.46</td>
<td>48.04</td>
<td>−4.55</td>
<td>136.84</td>
<td>0.00</td>
<td>0.77</td>
<td>[0.43, 1.21]</td>
</tr>
<tr>
<td>8</td>
<td>26.56</td>
<td>48.24</td>
<td>−4.79</td>
<td>136.39</td>
<td>0.00</td>
<td>0.81</td>
<td>[0.43, 1.21]</td>
</tr>
</tbody>
</table>

*a Welch’s t-tests were conducted.

*b A Bonferroni correction for multiple comparisons was applied.

**c** 95% CIs of effect size generated via bootstrapping (n = 10,000) with the bca method.

Confidence intervals. This indicated a small effect of condition on quiz performance Cohen’s $d = 0.28$, 95% CI $[-0.07, 0.62]$, which could explain why those in the high enthusiasm condition $(M = 0.70, SD = 0.18)$ performed nominally better on the quiz compared to those in the low enthusiasm condition $(M = 0.65, SD = 0.16)$; see Figure 3. However, it is critical to note that the values of 95% CI ranged from a very small negative effect to a moderate positive effect. Thus, this result did not provide conclusive evidence that instructor enthusiasm has any impact on memory performance.

Summary

Participants’ reports of instructor enthusiasm revealed that our manipulation was successful. More importantly, participants reported greater engagement when viewing the video with higher instructor enthusiasm than the one with lower enthusiasm, and the effect that was slightly smaller at the beginning of the video than during the rest of it. Furthermore, in each enthusiasm condition, participants’ ratings of their engagement increased as their ratings of instructor enthusiasm increased. Finally, our manipulation of instructor enthusiasm did not influence memory for the lecture material. Thus, instructor enthusiasm seems to influence subjective levels of engagement without concurrently having a substantial effect on learning.

Study 2

While in Study 1 we did not find an immediate benefit of instructor enthusiasm on memory for lecture content, it is possible that increasing instructor engagement may lead to other, less immediate, benefits for participants. For example, students who experience a more enthusiastic online lecture might be more...
inclined to study the course material outside of class time, or they might be more likely to access/virtually attend the next online lecture. More generally, there might be motivational benefits to experiencing an enthusiastic lecture and feeling more engaged (Horovitz and Mayer, 2021; Lawson et al., 2021). To explore this possibility, in Study 2 we followed the same procedures as Study 1, with the exception that at the end of the lecture participants answered a question assessing their motivation to watch the next lecture in the series. This question was placed prior to the quiz, so that participants’ performance and experience with the quiz did not influence their self-reports of motivation. We expected to replicate the main findings of Study 1, and also to find that instructor enthusiasm would have a positive impact on participant motivation, such that those participants who were in the enthusiastic condition would be more motivated to watch the next lecture in the series.

**Methods**

Our methods for Study 2 followed largely identical procedures and measures as Study 1, other than the addition of the motivation measure before the final quiz, and a new sample of participants. For brevity, we have elected not to repeat the methodology and only note the differences between the studies.

**Participants**

Participants were recruited from Psychology courses at the University of Waterloo in exchange for partial course credit. After following the same screening procedures as used in Study 1, 204 participants (139 female, 69 male, 2 non-binary or non-conforming, 1 prefers not to answer)—110 in the enthusiastic condition, and 94 in the non-enthusiastic condition—were included in the final sample.

**Measures**

For Professor Enthusiasm, Engagement, Multitasking, and Memory Performance, see Study 1.

**Motivation**

To assess how the enthusiasm manipulation impacted participant motivation to watch the next lecture, we added an additional question, presented following the end of the lecture and prior to the quiz. Specifically, we asked participants to “Imagine the lecture you just watched was part of one of your online courses. Based on your experiences in the lecture you just watched, how motivated would you be to watch the next lecture from the course?” Participants used a slider to rate their level of motivation on a scale from 0 (not at all motivated) to 100 (extremely motivated).

**Procedure**

In Study 2, we followed the same procedure as Study 1, with the addition of the motivation question added prior to the final quiz.

**Data cleaning**

In Study 2, we used the same data cleaning procedure as Study 1. Changes were made to accommodate the addition of the motivation question and to score the quiz. For additional details, see the R code for Study 2. The raw data and code used for Study 2 can be found on the OSF page: [https://osf.io/t9drh/](https://osf.io/t9drh/).

**Results and discussion**

As in Study 1, statistical analyses were conducted using R (R Core Team). Importantly, we conducted the same analyses that we conducted

---

15 Some minor changes and improvements were made for the sake of accommodating our additional variables (i.e., motivation), but these changes are related to data processing, not analyses.

16 In Study 2, we recorded the participants’ specific responses (i.e., A, B, C, D) in the raw data. In Study 1, these results were automatically scored as 1 (correct) or 0 (incorrect) when the data was sent to the server.
in Study 1, with the expectation that our findings would replicate. The ANOVAs were conducted using the ez package (Lawrence, 2016) and follow-up t-tests were conducted using the statsmodels package (Kassambara, 2023a,b). Further, bootstrapped 95% CIs of effect size using the bias-corrected and accelerated method (Efron, 1987; Kirby and Gerlanc, 2013) were generated using statsmodels (Kassambara, 2023a,b). The raw data and the full code used to clean and analyze the data for is available on the study OSF page: https://osf.io/t9drh/.

### Manipulation check

We started by analyzing ratings of instructor enthusiasm to determine whether the effect of manipulation replicated in Study 2. Once again, we conducted a 2×2×8 mixed ANOVA with Condition (High vs. Low Enthusiasm) and Version (A or B) as the between-subjects factors and Time (Probes 1 to 8) as the within-subjects factor. Mauchley’s test indicated that the assumption of sphericity was violated $W = 0.13$, $p < 0.05$, so a Greenhouse-Geiser correction was applied. Consistent with Study 1, the omnibus ANOVA indicated a significant main effect of Condition $F(1, 200) = 155.86$, $p < 0.001$, $\eta^2_p = 0.39$, indicating the manipulation was successful (see Table 5).

Unlike Study 1, there was also a significant main effect of Time (see Table 5), a significant interaction between Time and Version (see Table 5), and a significant 3-way interaction between Condition, Version and Time $F(3.88, 775.07) = 2.57$, $p = 0.039$, $\eta^2_p = 0.00$ (see Table 5). To follow up the highest-level interaction (i.e., the 3-way interaction), we split the data by Version (i.e., A or B), and conducted a mixed ANOVA with Condition (High vs. Low Enthusiasm) as the between-subjects factor and Time (Probes 1 to 8) as the within-subjects factor. The results of Version A are presented in Table 6, and the results of Version B are presented in Table 7.

As can be seen in the tables, there was no significant interaction between condition and probe number in either Version. In Figure 4, we present a visualization of the relation between Condition and Time, split by Version. Examining the figure also suggests that the effect of Condition is consistently large across Time, and does not notably differ as a function of Version. The only significant effect in either ANOVA was a significant effect of Time in Version A. A full table of the non-collapsed means can be found in Supplementary Tables on the associated OSF page.

Given that there was no significant interaction between Condition and Time in either Version A or B, we were able to return to our planned analyses and examined the bootstrapped effect sizes of condition at each probe (see Table 8). As can be seen in Table 8, the effect of Condition on ratings of instructor enthusiasm was consistently large across probes (i.e., there is a large effect at each probe). This result is consistent with Study 1, and once again indicates a successful manipulation of instructor enthusiasm.

### The effect of instructor enthusiasm on participant engagement

Next, to address our main question of interest, we examined the effect of enthusiasm on participants’ ratings of their engagement during the lecture. We conducted a 2×2×8 mixed ANOVA on engagement, with Condition (High vs. Low Enthusiasm) and Version (A or B) as the between-subjects factors, and Time (Probe 1 to 8) as the within-subjects factor. Mauchley’s test indicated the assumption of sphericity was violated $W = 0.10$, $p < 0.05$, so a Greenhouse-Geiser correction was applied. Consistent with Study 1, there was a significant main effect of Condition $F(1, 200) = 36.04$, $p = 0.001$, $\eta^2_p = 0.12$, such that participant engagement was substantially higher when participants watched the high enthusiasm compared to the low enthusiasm version of the lecture. There was no significant main effect of Version $F(1, 200) = 0.04$, $p = 0.850$, $\eta^2_p = 0.00$. As in Study 1, there was a significant main effect of Time $F(3.70, 739.78) = 22.54$, $p < 0.001$, $\eta^2_p = 0.02$. There was also a significant interaction between Version and Time $F(3.70, 739.78) = 3.85$, $p = 0.001$, $\eta^2_p = 0.03$.

---

**Table 5** Sample 2 ANOVA results for ratings of instructor enthusiasm.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$df_{Num}$</th>
<th>$df_{Den}$</th>
<th>Epsilon</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>1.00</td>
<td>200.00</td>
<td>155.86</td>
<td>0.000</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>1.00</td>
<td>200.00</td>
<td>0.60</td>
<td>0.438</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Condition × Version</td>
<td>1.00</td>
<td>200.00</td>
<td>0.06</td>
<td>0.805</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>Time (probe)</td>
<td>3.88</td>
<td>775.07</td>
<td>0.55</td>
<td>3.11</td>
<td>0.016</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition × Time (probe)</td>
<td>3.88</td>
<td>775.07</td>
<td>0.55</td>
<td>1.94</td>
<td>0.105</td>
<td>0.00</td>
</tr>
<tr>
<td>Version × Time (probe)</td>
<td>3.88</td>
<td>775.07</td>
<td>0.55</td>
<td>3.04</td>
<td>0.018</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition × Version × Time (probe)</td>
<td>3.88</td>
<td>775.07</td>
<td>0.55</td>
<td>2.57</td>
<td>0.039</td>
<td>0.00</td>
</tr>
</tbody>
</table>

$df_{Num}$ indicates degrees of freedom numerator; $df_{Den}$ indicates degrees of freedom denominator; Epsilon indicates Greenhouse-Geiser multiplier for degrees of freedom, $p$-values and degrees of freedom in the table incorporate this correction. $\eta^2_p$ indicates generalized eta-squared.

---

**Table 6** Sample 2 Anova results for ratings of instructor enthusiasm (Version A).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$df_{Num}$</th>
<th>$df_{Den}$</th>
<th>Epsilon</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>1.00</td>
<td>97.00</td>
<td>72.62</td>
<td>0.000</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>4.13</td>
<td>400.47</td>
<td>3.95</td>
<td>0.003</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>Condition × Version</td>
<td>4.13</td>
<td>400.47</td>
<td>0.59</td>
<td>2.13</td>
<td>0.074</td>
<td>0.00</td>
</tr>
</tbody>
</table>

$df_{Num}$ indicates degrees of freedom numerator; $df_{Den}$ indicates degrees of freedom denominator; Epsilon indicates Greenhouse-Geiser multiplier for degrees of freedom, $p$-values and degrees of freedom in the table incorporate this correction. $\eta^2_p$ indicates generalized eta-squared.

---

**Table 7** Sample 2 ANOVA results for ratings of instructor enthusiasm (Version B).

<table>
<thead>
<tr>
<th>Predictor</th>
<th>$df_{Num}$</th>
<th>$df_{Den}$</th>
<th>Epsilon</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2_p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>1.00</td>
<td>103.00</td>
<td>83.31</td>
<td>0.000</td>
<td>0.40</td>
<td></td>
</tr>
<tr>
<td>Version</td>
<td>3.39</td>
<td>348.89</td>
<td>0.48</td>
<td>2.01</td>
<td>0.104</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition × Version</td>
<td>3.39</td>
<td>348.89</td>
<td>0.48</td>
<td>2.39</td>
<td>0.061</td>
<td>0.00</td>
</tr>
</tbody>
</table>

$df_{Num}$ indicates degrees of freedom numerator; $df_{Den}$ indicates degrees of freedom denominator; Epsilon indicates Greenhouse-Geiser multiplier for degrees of freedom, $p$-values and degrees of freedom in the table incorporate this correction. $\eta^2_p$ indicates generalized eta-squared.
We did not investigate this interaction further, as it did not speak to our main question of interest (i.e., the effect of condition on engagement). No other effects were significant (see Table 9, see the OSF page for a full table of means).

To be consistent with Study 1 and our planned analyses, we once again examined the impact of instructor enthusiasm on engagement at each probe. To do so, we once again used Welch’s t-test with a Bonferroni correction for multiple comparisons and bootstrapped 95% CIs of the effect size, using the rstatix package (Kassambara, 2023a,b).

As can be seen in Table 10, the effect of Condition was significant at each probe, such that those in the high enthusiasm condition were significantly more engaged than those in the low enthusiasm condition.

TABLE 8 Sample 2 descriptive statistics, post-hoc tests and 95% CIs for ratings of instructor enthusiasm.

<table>
<thead>
<tr>
<th>Time</th>
<th>Low enthusiasm</th>
<th>High enthusiasm</th>
<th>t^a</th>
<th>df^a</th>
<th>p.adj^b</th>
<th>d</th>
<th>95% CI^c</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>22.79</td>
<td>54.70</td>
<td>-11.18</td>
<td>201.64</td>
<td>0.00</td>
<td>1.56</td>
<td>[1.21, 1.89]</td>
</tr>
<tr>
<td>2</td>
<td>23.41</td>
<td>58.36</td>
<td>-11.98</td>
<td>201.58</td>
<td>0.00</td>
<td>1.68</td>
<td>[1.33, 2.03]</td>
</tr>
<tr>
<td>3</td>
<td>21.52</td>
<td>56.40</td>
<td>-11.99</td>
<td>200.84</td>
<td>0.00</td>
<td>1.67</td>
<td>[1.32, 2.03]</td>
</tr>
<tr>
<td>4</td>
<td>21.76</td>
<td>56.85</td>
<td>-11.10</td>
<td>200.82</td>
<td>0.00</td>
<td>1.54</td>
<td>[1.19, 1.91]</td>
</tr>
<tr>
<td>5</td>
<td>21.19</td>
<td>55.75</td>
<td>-11.18</td>
<td>201.66</td>
<td>0.00</td>
<td>1.56</td>
<td>[1.21, 1.91]</td>
</tr>
<tr>
<td>6</td>
<td>20.09</td>
<td>53.34</td>
<td>-10.56</td>
<td>199.64</td>
<td>0.00</td>
<td>1.47</td>
<td>[1.13, 1.82]</td>
</tr>
<tr>
<td>7</td>
<td>19.53</td>
<td>57.56</td>
<td>-12.05</td>
<td>196.71</td>
<td>0.00</td>
<td>1.67</td>
<td>[1.33, 2.03]</td>
</tr>
<tr>
<td>8</td>
<td>18.39</td>
<td>56.02</td>
<td>-11.81</td>
<td>193.68</td>
<td>0.00</td>
<td>1.64</td>
<td>[1.29, 1.99]</td>
</tr>
</tbody>
</table>

^aWelch’s t-tests were conducted.
^bA Bonferroni correction for multiple comparisons was applied.
^c95% CIs of effect size generated via bootstrapping (n = 10,000) with the bca method.

TABLE 9 Sample 2 ANOVA results for ratings of engagement.

<table>
<thead>
<tr>
<th>Predictor</th>
<th>df^Num</th>
<th>df^Den</th>
<th>Epsilon</th>
<th>F</th>
<th>p</th>
<th>η^2_g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Condition</td>
<td>1.00</td>
<td>200.00</td>
<td></td>
<td>36.04</td>
<td>0.000</td>
<td>0.12</td>
</tr>
<tr>
<td>Version</td>
<td>1.00</td>
<td>200.00</td>
<td></td>
<td>0.04</td>
<td>0.850</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition × Version</td>
<td>1.00</td>
<td>200.00</td>
<td></td>
<td>0.18</td>
<td>0.674</td>
<td>0.00</td>
</tr>
<tr>
<td>Time (probe)</td>
<td>3.70</td>
<td>739.78</td>
<td>0.53</td>
<td>22.54</td>
<td>0.000</td>
<td>0.02</td>
</tr>
<tr>
<td>Condition × Time (probe)</td>
<td>3.70</td>
<td>739.78</td>
<td>0.53</td>
<td>1.78</td>
<td>0.137</td>
<td>0.00</td>
</tr>
<tr>
<td>Version × Time (probe)</td>
<td>3.70</td>
<td>739.78</td>
<td>0.53</td>
<td>3.85</td>
<td>0.005</td>
<td>0.00</td>
</tr>
<tr>
<td>Condition × Version × Time (probe)</td>
<td>3.70</td>
<td>739.78</td>
<td>0.53</td>
<td>1.28</td>
<td>0.277</td>
<td>0.00</td>
</tr>
</tbody>
</table>

df^Num indicates degrees of freedom numerator, df^Den indicates degrees of freedom denominator. Epsilon indicates Greenhouse–Geisser multiplier for degrees of freedom, p-values and degrees of freedom in the table incorporate this correction. η^2_g indicates generalized eta-squared.
Importantly, the effect of enthusiasm once again ranged from moderate to large at each time point—indicating an effect of enthusiasm on engagement throughout the lecture. Thus, these results replicate our results from Study 1, indicating that those in the high enthusiasm condition were consistently more engaged during the lecture than those in the low enthusiasm condition, and they provide support for the notion that instructor enthusiasm impacts student engagement.

The relation between ratings of enthusiasm and engagement

Next, we examined the relation between participants’ perception of instructor enthusiasm and engagement using correlational analyses. As in Study 1, there was a significant effect of enthusiasm condition on both enthusiasm and engagement ratings, so we computed a separate correlation for each condition. Consistent with Study 1, there was a positive relation between enthusiasm ratings and engagement in both conditions, such that those who found the professor to be more enthusiastic/energetic tended to indicate they were more engaged during the lecture. The correlation was nominally higher in the low enthusiasm condition \( r(92) = 0.56, 95\% \text{ CI} [0.40 \text{ to } 0.68], \) than the high enthusiasm condition \( r(108) = 0.40, 95\% \text{ CI} [0.22 \text{ to } 0.54], \) however, the confidence intervals for the correlations were overlapping, suggesting the correlation was once again of a similar magnitude in both conditions. The pattern we observed in Study 1 (i.e., where the positive relation appeared to tail off at the top of the enthusiasm ratings), was not clearly present in Study 2 (see Figure 5). Critically, in general, these results replicate the pattern we observed in Study 1, indicating a positive relation between instructor enthusiasm/energy and engagement.

The effect of instructor enthusiasm on quiz performance

As in Study 1, we once again investigated whether the effect of enthusiasm impacted participants’ memory for lecture content as assessed using a multiple-choice quiz. We conducted a 2×2 ANOVA with Condition (High vs. Low Enthusiasm) and Version (A or B) as between-subject factors. Consistent with Study 1, there was no significant main effect of Condition \( F(1, 200) = 1.95, p = 0.165, \eta^2_p = 0.01, \) Version \( F(1, 200) = 0.66, p = 0.419, \eta^2_p = 0.00, \) or the interaction \( F(1, 200) = 3.63, p = 0.058, \eta^2_p = 0.02. \) As before, we examined the effect size of the difference between the two conditions by collapsing across Version, and generating a bootstrapped confidence interval. This indicated a small effect of

Condition on quiz performance Cohen’s \( d = 0.196 95\% \text{ CI} [-0.08, 0.47], \) however, the 95% CIs included both negative and positive estimates, suggesting the effect is negligible (see Figure 6). Thus, as in Study 1, those in the high enthusiasm condition \( (M = 0.66, SD = 0.19) \) performed only nominally better on the quiz than those in the low enthusiasm condition \( (M = 0.62, SD = 0.17), \) meaning we did not find any evidence to support an effect of instructor enthusiasm on memory performance. 

Motivation

In Study 2, we addressed an additional research question examining the effect of instructor enthusiasm on participants’ motivation—with the expectation that those who viewed an enthusiastic lecture would be more motivated to watch the next lecture from the course. To investigate this question, we conducted a 2×2 ANOVA on motivation with Condition (High vs. Low Enthusiasm) and Version (A or B) as between-subject factors. There was a significant main effect of condition...
Marty-Dugas et al.

$F(1, 200) = 47.51, p < 0.001, \eta^2_g = 0.19$, such that those in the high enthusiasm condition ($M = 49.18, SD = 27.16$) reported being significantly more motivated to watch the next lecture than those in the low enthusiasm condition ($M = 24.49, SD = 23.50$; see Figure 7) \(^{19}\). Bootstrapped confidence intervals indicated that this effect ranged from moderate to large $\text{Cohen's } d = 0.97$, 95% CI [0.65, 1.29]. There was no significant main effect of Version $F(1, 200) = 1.92, p = 0.167, \eta^2_g = 0.01$, nor was there a significant interaction Version $F(1, 200) = 0.05, p = 0.827, \eta^2_g = 0.00$. Thus, these results seem to suggest that instructor enthusiasm has a noteworthy effect on participants' motivation to engage with further lecture content in a course.

\(^{19}\) The full table of means split by Version can be found in Supplementary Tables on the associated OSF page.

Summary

Broadly, the results of Study 2 replicated the main findings of Study 1. Participant reports of instructor enthusiasm once again indicated that our manipulation was successful. Further, participants once again reported substantially greater engagement when viewing the high enthusiasm lecture, an effect that 95% CIs indicated was moderate to large across the length of the lecture. Also consistent with Study 1, we found that in both conditions, participants’ ratings of engagement increased as their ratings of instructor enthusiasm increased, and that instructor enthusiasm did not influence memory for lecture material, as assessed using an immediate multiple-choice quiz.

In addition, in Study 2 we examined a new question concerning the influence of instructor enthusiasm on participant motivation by asking participants to imagine the lecture they had just watched was part of one of their online courses. There was a moderate to large effect of instructor enthusiasm, such that participants were substantially more motivated to watch next lecture from the course when they had viewed the enthusiastic lecture.

General discussion

In two studies, we investigated the impact of instructor enthusiasm on student engagement and memory performance in the context of an online, asynchronous lecture. Participants were randomly assigned to either a high enthusiasm or low enthusiasm condition, in which the instructor delivered the lecture either with, or without, enthusiasm. Importantly, all participants were presented with the same lecture (i.e., on sleep and circadian rhythms), delivered by the same instructor, using the same slides, and the same verbal script. The only difference between the conditions was the instructor’s vocal delivery of the material, which was either higher or lower in enthusiasm. Across both Study 1 and Study 2 our analyses indicated that there was a large effect of instructor enthusiasm on participants’ ratings of instructor enthusiasm, which persisted across the length of the lecture, indicating that our manipulation of enthusiasm was successful and enduring.

Critically, regarding our primary aim, both studies showed that there was a consistent effect of enthusiasm on participants’ level of engagement during the lecture, such that those in the high enthusiasm condition were significantly more engaged than those in the low enthusiasm condition. While in Study 1 the effect of enthusiasm appeared to be smaller early in the lecture and consistently larger over the rest of the lecture, this pattern did not replicate in Study 2; instead, the effect was a consistent magnitude across the length of the lecture (i.e., the 95% CIs of effect size ranged from small to large at each time point). Most importantly, the results of both studies provide clear evidence that there was an effect of enthusiasm on participant engagement—such that those in the high enthusiasm condition were more engaged throughout the lecture, compared to those in the low enthusiasm condition.

In addition, we conducted correlational analyses to examine the relation between instructor enthusiasm and participant engagement at the level of individual differences. Because of the significant effect of condition on both ratings of enthusiasm and engagement, we conducted these analyses split by condition. In both studies, we found a moderate positive correlation between enthusiasm and engagement, in both the high and low enthusiasm conditions. In other words, the more enthusiastic participants found the instructor, the more engaged they...
felt regardless of the overall level of enthusiasm the instructor displayed (i.e., regardless of which condition they were in). Interestingly, in Study 1, graphical analysis suggested that the positive nature of the relation appeared to tail off and decline at the uppermost levels of enthusiasm. However, this pattern was not clearly present in Study 2. Thus, while it is possible that extremely high levels of enthusiasm may be detrimental or off-putting to some students (McKinney et al., 1983), we did not observe consistent evidence of such a pattern in the present study. More importantly, however, is the consistent positive relation across both conditions, and across Study 1 and Study 2.

We also examined the impact of instructor enthusiasm on participants memory for lecture content, as assessed by their performance on an immediate multiple-choice quiz. While participants in the high enthusiasm condition had nominally higher performance on the quiz in Study 1, this effect was not significant. Consistent with these results, there was no evidence for a significant effect of enthusiasm on memory performance in Study 2. Further, in both studies the range of possible values identified by the bootstrapped confidence intervals indicated that the pattern of data was just as consistent with a small negative effect as with a moderate positive effect. As such, given the current findings, it is not possible to determine whether the effect of enthusiasm on memory performance is positive, negative, or non-existent, limiting the conclusions that can be drawn about the effect of enthusiasm on memory in this study.

Why might we have failed to find an influence of instructor enthusiasm on memory performance even though it did influence reports of engagement? One possible explanation is that the study may have been limited by the use of self-reports to assess engagement, and that these reports do not reflect people's actual engagement with the video content. While this may be true, many prior studies have established a link between subjective reports of attentional engagement and performance, suggesting that subjective reports can be reliable and valid (Cheyne et al., 2009; Smilk et al., 2010; Risko et al., 2012, 2013; Szpunar et al., 2013a; Kane et al., 2017; Wammes et al., 2019; Marty-Dugas et al., 2021). Another possibility is that the memory test was not sensitive enough to detect engagement-related differences between conditions. In the present study memory was assessed using a recognition test after a short delay, which may have only tapped a short-term shallow level of encoding, thus limiting the conclusions we can draw about memory more generally. That is, our findings do not rule out the possibility that enthusiasm may impact memory performance at deeper levels of processing, or when there is a delay between encoding and testing (see Roediger and Karpicke, 2006). Further, it could also be the case that the lecture was of a high enough quality in terms of its other characteristics (e.g., quality of the slides and visualizations), that participants were able to learn the material sufficiently even in the low enthusiasm condition. These various possibilities could be addressed in future studies.

Finally, although we did not find that instructor enthusiasm influenced memory for course content, we did find that instructor enthusiasm affected participants' intentions to watch another video on the same topic. In Study 2 we asked participants to imagine that the lecture they had just watched was part of an online course they were taking, and, based on their experience during the lecture, to rate how motivated they would be to watch the next lecture in the course. In line with our predictions, the results showed that those in the high enthusiasm condition clearly indicated they were more motivated to watch the next lecture. This finding could be particularly applicable to asynchronous courses, wherein students decide for themselves when the next lecture will be viewed. If higher instructor enthusiasm during online lectures can lead to higher motivation to watch the next lecture, it is possible that high enthusiasm may lead to students “attending” virtual lectures at more regular intervals. However, it is important to note that the present study was limited by the use of the single lecture, and measured student motivation by asking them to imagine their intention to watch the next lecture, rather than assessing their behavior. While intention is an important indicator of future behavior (Ajzen, 1991), investigating how instructor enthusiasm (in an online lecture) impacts behaviors that require students to re-engage with the material (watching the next lecture, re-watching lectures, attending office hours) would be a meaningful improvement for future work.

Educational implications

When considering the educational implications of the present findings, it is important to keep in mind the structure of the course being taught. For many university level courses, online lectures are delivered asynchronously using a pre-recorded video. It is these sorts of courses and lessons that the results of the present study are most directly applicable to, because the course structure is the most similar to the procedure used in the study. Put differently, any educational implications become more speculative as the structure of the course becomes more dissimilar from the procedure used in the present study. For example, a lesson being taught in-person to young elementary schoolchildren would be more dissimilar than an in-person lecture attended by university students. The extent to which these results generalize to different courses and student populations is a question for future research. Regardless, the results of the present study clearly indicated higher attention and motivation among the participants who viewed the high enthusiasm lecture—supporting the notion that increased teacher enthusiasm may be a promising as a strategy to capture student attention, and to motivate them to make a greater effort in their course.

Conclusion

In conclusion, despite the lack of an impact on memory performance, the present studies illustrate that varying instructor enthusiasm can influence students’ attentional engagement with a video lecture, as well as their motivation to engage with the course material in the future. These results are quite striking when one considers that in the present study instructor enthusiasm was modulated by changing the voice only. It is possible that over multiple lectures the effect of an enthusiastic instructor may wear off, but on the other hand, it is also possible that the effect might cascade over time and grow stronger with each lecture. Regardless, our findings suggest that increased instructor enthusiasm may be a promising strategy to improve student attention and motivation in a learning context.

---

20 Of course, ruling the effect out on the basis of the present result would mean making the error of accepting the null. However, we wish to illustrate why the present results cannot be used to draw this conclusion, even if one were to incorrectly accept the null.
Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://osf.io/9drh/.

Ethics statement

The studies involving humans were approved by the University of Waterloo Office of Research Ethics and the McMaster Research Ethics Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JMD: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MR: Conceptualization, Investigation, Methodology, Software, Visualization, Writing – original draft. RM: Project administration, Resources, Software, Writing – review & editing, JK: Funding acquisition, Writing – review & editing, DS: Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by an NSERC Discovery Grant awarded to DS, funding by the McCall MacBain Foundation provided to JK, and by the McCall MacRae Postdoctoral Fellowship in Education and Cognition awarded to JMD. These funding sources had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript the following grant number can be added for DS NSERC Discovery grant (2019-04071).

Acknowledgments

All data, as well as R code used to conduct the analysis have been made publicly available at the Open Science Foundation (OSF) and can be accessed at https://osf.io/9drh/.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References


R Core Team (2023). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. [Computer software]. Available at: https://www.R-project.org/


Ware, J. E., and Williams, R. G. (1975). The Dr. fox effect: a study of lecturer effectiveness and ratings of instruction. Acad. Med. 50, 149–156. doi: 10.1097/00001888-197502000-00006


