# The Role of Graphics in Video Lectures

by

Laura J. Bianchi

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2019

**Author's Declaration**

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

## Abstract

With the increase in online course use (Allen & Seaman, 2017), there is an increasing need to determine the most effective (i.e., the most conducive for learning) way to present lectures online (e.g., video lectures). Lecture graphics that are interesting but extraneous to the content (e.g., a celebrity), have been shown to impair comprehension of the material, likely resulting from an increase in cognitive load. In this study, the use of graphics on the slides of an online psychology lecture was manipulated to determine the extent to which images can improve (or impair) comprehension as well as the effect it may have on intentional and unintentional mind-wandering. Across our two experiments, we demonstrate no differences across conditions (i.e., unnecessary graphics, relevant graphics, no graphics) in overall comprehension and limited differences in mind wandering behaviour.

**Acknowledgements**

I would like to thank my supervisor Dr. Evan Risko for the support and guidance he gave me during this process. His enthusiasm and insight for cognitive research has inspired me to be a better researcher by learning and applying new methodologies throughout this degree program.

I would like to thank my readers, Dr. Derek Besner, and Dr. Jonathan Fugelsang, for their helpful comments and expertise. I would also like to thank Dr. Kristen Wilson for her help with the project's initial design, including the creation the stimuli used.

Lastly, I would like to thank my family and friends who helped me through every challenge that I faced, as well as my colleagues in the Cognition and Natural Behaviour Laboratory, for their friendship and support.

**Table of Contents**

**List of Tables**

**List of Figures**

# Chapter 1: General Introduction

## *1.1 Online learning and multimedia learning*

There has been a growing use of online courses in post-secondary education in the last 20 years. Colleges in the United States reported an increase in the proportion of students enrolled in an online course from 10% in 2002 to 30% in 2015 (Allen & Seaman, 2017). In Canada, the same pattern of results is apparent with 2/3 of post-secondary institutions reporting an increase in online course enrollment from the 2015-16 school year to the 2017-18 school year, half of which report an increase of 10% or more (Canadian Digital Learning Research Association, 2019). Given the rising use, there is a growing need for researchers to study these environments and determine ways to optimize learning.

Online learning environments rely heavily on video lectures as a means of communicating course content (e.g., Gorissen, van Bruggen & Jochems, 2012). Video lectures can include audio, images and text making them a form of multimedia. Research investigating multimedia learning has a long history (Mayer, Heiser & Lonn, 2001; Mayer, 1997; Mayer & Moreno, 2003; Najjar, 1996; Moreno & Mayer, 1999) and has clearly demonstrated that the decisions one makes about the combinations of media and their structure can have notable effects on learning (Mayer, 2005; Mayer, 2009). For example, Mayer and Anderson (1991) found that when presenting learners with audio of how a bike pump works, followed by an animation showing how the pump works, learners had poorer learning outcomes compared to those who were presented the audio and animation simultaneously. In the present investigation we continue this general line of inquiry by examining, in a postsecondary video lecture, the influence of the presence of graphics in lecture slides and in particular the influence of the graphic's relevance to the lecture content on comprehension and attentional engagement with the lecture material.

*1.2 Cognitive load theory and unnecessary graphics*

The dominant framework for understanding multimedia learning is cognitive load theory (Mayer, 2009; Sweller, 2010). This theory assumes that learners have a limited capacity for information processing, and therefore can experience overload, thus impairing learning (Mayer, 2005; Mayer, 2009; Sweller, 2010). Consequently, decisions about what forms of multimedia to use need to consider how it impacts the load imposed on the learner.

One common combination of media used in a university lecture is merging audio or text with graphics. Research investigating the use of graphics on comprehension has found that in some cases (e.g., when the graphics are unnecessary), the addition of graphics can have a negative impact (Harp & Mayer, 1997; Mayer, Griffith, Jurkowitz & Rothman, 2008; Sanchez & Wiley, 2006; Sung & Mayer, 2012). For example, Harp and Mayer (1997) conducted a series of experiments comparing the addition of "seductive" graphics (i.e., irrelevant to learning but catchy) to conditions without those graphics and found that comprehension was reduced in the former condition even though it increased ratings of interest.

This negative effect on comprehension due to unnecessary graphics has been found across various studies (Garner, Gillingham & White, 1989; Harp & Maslich, 2005; Park, Moreno, Seufert & Brunken, 2010; Sanchez & Wiley 2006), and is thought to be caused by distracting the learner (Harp & Mayer, 1998) thus, consuming some of the limited resources potentially available to learn the material (Mayer, 2009; Sung & Mayer, 2012; Paas, Renkl & Sweller, 2003). That said, the inclusion of unnecessary graphics do not always produce negative effects (Ketzer-Nöltge, Schweppe & Rummer, 2018; Kühl, Moersdorf, Römer & Münzer, 2018; Strobel, Grund & Lindner, 2018), and under low load conditions they may actually improve learning outcomes (Park et al., 2010). These inconsistences in the literature have led some researchers to believe that the negative effect of unnecessary graphics may have been exaggerated, since the effect is often small and can be moderated by a various other factors (e.g., prior knowledge; Eitel & Kuhl, 2018).

*1.3 Mind wandering and unnecessary graphics*

While the primary consideration in designing video lectures is comprehension, comprehension is intimately tied up with the video lecture's ability to maintain learner's attentional engagement (Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012). Indeed, one can consider the capacity limit noted above to reflect our attentional capacity limits. That is, if a learner is attending to something other than the lecture, then the learner is not devoting attention to the lecture content, impairing their ability to store in memory and comprehend the material (McVay & Kane, 2012; Unsworth & McMillan, 2013; Risko et al., 2012; Smallwood & Schooler 2006). This consideration is arguably even more important when learning is online because individuals are engaging with the material on their own schedule (Gorissen et al., 2012), where they can easily disengage through media multi-tasking (e.g., switching to social media; Burak, 2012; Ralph, Thomson, Cheyne & Smilek, 2014), or quit the lecture early (Kim et al., 2014).

A popular means of indexing the extent to which an individual is attending to the lecture is to measure mind wandering – the shifting of attention from a task to internal thoughts (Smallwood & Schooler, 2006). Recent studies have suggested that mind wandering occurs often in video lectures (approximately 40% of the time; Risko et al., 2012), which can lead to impaired comprehension of the lecture material (Farley, Risko & Kingstone, 2013; Risko et al., 2012; Risko, Buchanan, Medimorec & Kingstone, 2013). Given that a lack of attentional engagement in a lecture will have negative consequences on comprehension, it is important for researchers to investigate ways to keep learners' attention on task with their lecture designs. For example, some studies suggest that time spent in a lecture can change rates of mind wandering, such that as the time into the lecture increases (i.e., the longer the learner is required to sustain their attention to the task) mind wandering increases (Farley et al., 2013; Risko et al., 2012; Szpunar, Multon & Schater, 2013).

In addition to time-on-task, research has suggested that some combinations of media (e.g., audio, text and graphics) may lead to more or less motivation for learners to continue to engage with the learning material (Moreno & Mayer 2007). Specifically, the use of attractive illustrations in a lecture have been demonstrated to be helpful (Magner, Schwonke, Aleven, Popescu, & Renkl 2014), and can increase interest in study material (Harp & Mayer, 1997), even though in the latter case these images impaired comprehension. If it is the case that unnecessary graphics keep learners engaged and interested in the material, then mind wandering might decrease in conditions where these types of graphics are used. Nonetheless, while attractive graphics may keep attention to the lecture through reduced mind wandering, they may still impair (or not benefit) comprehension if they are irrelevant to the lecture material. That is, unnecessary graphics could draw the learner's attention to an irrelevant aspect of the lecture (the graphic) and thus while their attention might be "on task" resources are being diverted from the to-be-learned material (Harp & Mayer 1998). Thus, a deeper understanding of how graphics influence attention, as indexed by mind wandering, would be valuable in informing instructional design.

*1.4. Present Investigation*

In the present investigation we examined the use of unnecessary graphics in an online video lecture, specifically, how it affects participant's comprehension and mind wandering during the lecture. Participants viewed one of three types of video lecture; 1) unnecessary graphics, 2) relevant graphics, and 3) no graphics, and completed a comprehension test to measure learning. In addition, during the lecture participants were probed to determine if they were mind wandering and if so what type (i.e., intentional or unintentional). Intentional mind wandering is considered to be a purposeful shift from the task to an internal thought, whereas unintentional mind wandering is considered shifting attention away from the task even though the intent was to remain on task (Seli, Carriere & Smilek, 2015; Seli, Risko & Smilek, 2016; Seli, Risko, Smilek & Schacter, 2016). Provided these types of mind wandering are argued to be distinct, measuring both provides a fuller understanding of how certain lecture designs can affect attentional engagement. In addition to mind wandering, we include an exploratory measure, with the intention of providing further insight into disengagement from the lecture. Specifically, we include a behavioural measure of "distraction" that consisted of recording how often participants clicked away from the task screen showing the lecture (i.e., our participants were completing the task online, as opposed to in a laboratory, and as such could freely disengage in this manner). This provided a behavioral measure of disengagement that complemented the self-report based mind wandering measure.

We were also interested in participant's perceptions of how much they learned from the lecture and of the graphics used. As such, participants were asked to provide judgments of learning (Wilson, Martinez, Mills, D'Mello, Smilek & Risko, 2018) before the comprehension test, and concluded the study with a rating scale of how helpful and relevant the graphics used were. The latter served as a manipulation check to determine if participants who were in the conditions with graphics, perceived the graphics as we intended them.

## Chapter 2: Examining Graphics in Video Lectures

### 2.1 Experiment 1

#### *2.1.1Methods*

*Participants.* A total of 215 people participated in this online study through Amazon's Mechanical Turk. Participants received $5 in compensation for completion of the study. An a-priori sample size was determined in order to achieve 0.80 power, at alpha .05, to detect a small-to-medium effect size of $\eta^2 = 0.05$ in a one-way omnibus ANOVA.

*Stimuli.* The video lecture consisted of a 25-minute slideshow presentation paired with audio of the instructor explaining the content of the lecture (i.e., an introduction to reasoning and decision making). The audio of the lecture was recorded before creating the matching slideshow presentation and, participants in all conditions listened to the same audio track. The lecture slides were developed using the audio recording. The text and positioning of the text on all of the lecture slides were matched across conditions, however the images that accompanied the text on the slide varied by condition (see Figure 1). For both conditions containing graphics, there was at minimum one graphic per slide, some contained up to 4 images. In all conditions, there was 71 slides which made up the presentation. In the unnecessary graphics condition, images that had no additional explanatory value with respect to the material but were engaging (e.g., cat photos) were featured on each slide of the presentation. In the relevant graphics condition, there were images that related to the content of the lecture (e.g., exemplars, infographics) presented on each slide. In the no graphics condition, there were no images of any kind on the slides.

*Measures.* Participants completed a demographic questionnaire where they reported their age, gender, highest level of education, current student status, as well as the number of online courses and the number of psychology courses they had taken. They were then presented with two pre-assessment questions assessing their existing knowledge of the content for the online lecture (i.e., reasoning and decision making). These questions were open response and participants were told they could respond with “I don’t know” if they did not know the answer and did not want to guess (see Appendix A).

During the video lecture participants received 10 mind wandering probes. The probe would appear at the same time for everyone, however they were presented at unequal intervals to prevent participants from anticipating the probe. The unequal intervals were decided using a random number generator, which lead to 6 of the probes occurring in the first half of the video (i.e., the first 12.5 mins) and the remaining 4 were distributed in the last half of the video (see Appendix B for timings). The question asked if the participant had been mind wandering in the moments before the probe and participants needed to select one of the following three responses: “yes -I was intentionally mind wandering”, “yes-I was unintentionally mind wandering”, or “no-I was fully focused on the lecture”.

After the video participants provided an estimate of the grade they expected on the upcoming test (judgement of learning). The study concluded with a test of the content of the lecture they just watched. There were 15 comprehension questions (10 multiple choice, 5 open response, see Appendix A). These test questions were additionally categorized as being “deep” (i.e., required integration/ manipulation of lecture material), or “shallow” (i.e., vocabulary, definitions, or have content verbatim from the lecture). Three of the test questions were categorized as deep learning questions, and the remaining 12 were categorized a shallow learning questions.

As a manipulation check, we asked participants at the end of the study to rate on a 1-7 scale how helpful they thought the graphics in the lecture were, and how relevant they thought the graphics were, where 1 meant completely relevant/helpful and 7 meant completely irrelevant/unhelpful. For easier data interpretation, we reverse scored responses so that a larger number represented greater helpfulness and relevance.

Additionally, throughout the study, if the participant navigated away from the task screen (e.g., minimizing the page featuring our study) and the page with our study was no longer in focus, the change was coded and timestamped. Participants were not told at the beginning of the study that this data would be recorded. We included this exploratory variable as a behavioural measure of "off-task" performance.

*Procedure.* After providing consent, participants were asked to complete the demographic questions and the pre-assessment questions. After the pre-assessment, participants were instructed to read a short description about what mind wandering is and what the different types of mind wandering are (i.e., intentional and unintentional). Participants were asked a comprehension (multiple choice) question about the difference between intentional and unintentional mind wandering to ensure they knew how to respond when probed during the video. The participants were then instructed to watch a 25-minute psychology lecture on reasoning and decision making, where they responded to the mind wandering probes. Participants knew they would be asked questions about the video. After the video participants completed the judgement of learning (JOL) and the comprehension test. Lastly, they were asked how helpful and how relevant they thought the graphics were in the lecture they watched. Once these questions had been completed participants were debriefed and compensated.

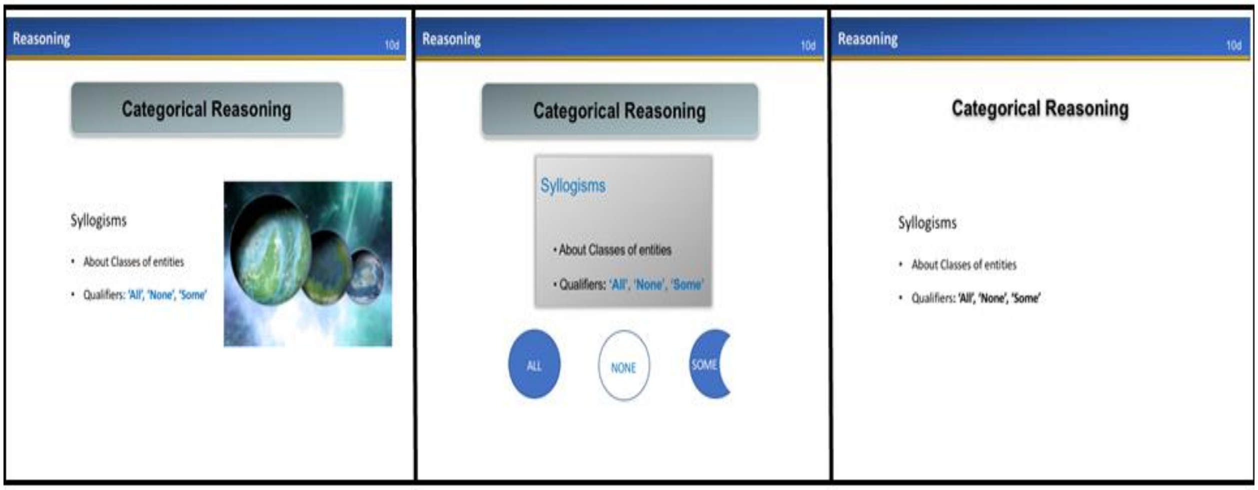


*Figure 1:* An example slide taken from each of the three lecture conditions: unnecessary graphics (left panel), relevant graphics (middle panel), no graphics (right panel).

*2.1.2 Results*

Twenty-nine participants were not included in data analysis since they did not complete all necessary aspects of the study (i.e., they quit before the post-assessment questions were completed). A further, 11 participants were excluded for failing to correctly answer the comprehension question on the difference between intentional and unintentional mind wandering. The distribution of the remaining 175 participants for each condition were as follows: 61 participants saw unnecessary graphics, 55 saw relevant graphics, and 59 saw no graphics. Demographic data was collected for all the participants. The mean age was 35 years old. There were 84 males and 91 females. Participants reported their highest level of education with 63 participants reporting having a high school diploma, 28 having a college diploma, 64 having a bachelor’s degree, 18 having a master’s degree, and 2 having a doctorate degree. A final question asked was if participants were currently university students, 146 reported not being a current student, and the remaining 29 said they were.

Open response comprehension questions (pre and post assessment questions) were scored by a research assistant who was blind to the condition participants were in. Participants could receive one mark for a correct response, half a mark for a partially correct response or zero for an incorrect response, therefore each question was scored out of 1. All correct multiple choice responses received a score of 1. The questions were then split by when the content was delivered in the video lecture, specifically if it fell within the first half of the video or the second half. Below we first report the effect of condition on accuracy in the test. Additionally, we split the questions by whether they can be considered deep or shallow learning questions, and question format (i.e., multiple choice or open responses). We then examine the effect of condition on the proportion of wandering events across the two types of mind wandering. Mind wandering was also analyzed across time (i.e., split by if the probe occurred at the start or end of the video). Lastly, we report an exploratory analysis of our “off-task” behavioural measure (i.e., deviation from the task screen). We concluded with an analysis for reports of helpfulness and relevance for the lectures which contained graphics, and conducted bivariate correlations for all of our dependent variables. We include a Bayes factor analysis for each effect using the default prior (0.71).

*Comprehension*. There was a marginally significant effect of condition on the proportion of correct answers for the pre-assessment questions, $F(2,172) = 2.91$, $p = .057$, $\eta_p^2 = .03$, $BF_{01} = 1.41$ (see Table 1), such that those in the relevant graphic condition had a greater proportion of correct answers compared to those in the no graphics condition, $t(112) = 2.45$, $p = .015$, $d = 0.46$, $BF_{10} = 2.87$. There was no difference between those in the relevant graphics and unnecessary graphics conditions, $t(114) = 0.87$, $p = .387$, $d = 0.16$, $BF_{10} = 0.27$, and no difference between those in the unnecessary graphics and those in the no graphics conditions, $t(118) = 1.64$, $p =$

.104, $d = 0.30$, $BF_{10} = 0.65$. Overall, the pre-assessment means were all low (i.e., below 15% correct) and the majority of participants (80%) had zero correct responses on the pre-assessment. Bayes factor analysis suggests that this marginal effect has slightly more support for the null.

There was no significant effect of condition on the proportion of correct answers to the comprehension test questions, $F(2,172) = 0.66$, $p = .518$, $\eta_p^2 < .01$, $BF_{01} = 9.77$. There was also no significant effect of condition on shallow learning questions, $F(2, 172) = 0.97$, $p = .380$, $\eta_p^2 < .01$, $BF_{01} = 7.47$ or deep learning questions, $F(2, 172) = 0.17$, $p = .843$, $\eta_p^2 < .01$, $BF_{01} = 14.90$. When we compared across the different question formats, we also found no effect of condition for either multiple choice, $F(2, 172) = 1.24$, $p = .293$, $\eta_p^2 = .01$, $BF_{01} = 5.96$ or open response questions, $F(2, 172) = 0.04$, $p = .959$, $\eta_p^2 < .01$, $BF_{01} = 16.65$ (see Table 1).

As noted above, there was a small difference across conditions in pre-knowledge of the lecture material and as such, we conducted a linear regression examining test scores across the conditions, controlling for pre-assessment accuracy. There was still no effect of condition on comprehension when we controlled for pre-knowledge, $F(2, 171) = 0.42$, $p = .661$.

To determine if there was an effect of time on task and if it differed by video condition, we split questions by whether the content occurred during the first half (first 12.5 mins) or second half of the video, and conducted a condition by time (first half, second half) mixed ANOVA. There was no significant main effect of time, $F(1, 172) = 2.52$, $p = .114$, $\eta_p^2 = .01$, $BF_{01} = 2.64$ and no interaction of time with condition, $F(2, 172) = 0.12$, $p = .885$, $\eta_p^2 < .01$, $BF_{01} = 15.79$ (see Table 1).

*Judgements of Learning.* Three participants did not respond to this question and therefore are not included in the analysis. There was no significant effect of condition for JOLs, $F(2, 169)$

= 0.70, $p$ = .499, $\eta_p^2$ < .01, $BF_{01}$ = 9.34 (see Table 1). When we compared actual test accuracy with JOLs using a paired t-test, there was a significant difference, $t(171) = 4.14$, $p < .001$, $d$ = 0.31, $BF_{10}$ = 262.22, such that participants reported higher JOLs ($M$ = 0.58) than their actual test score ($M$ = 0.51; i.e., they were overconfident).

*Table 1: Means and 95% CIs for the proportion of correct answers to the pre-assessment questions, the test questions, the test questions split by type (shallow, deep), question format (multiple choice, short answer) and time (start, end), and participants' JOL (score out of 100) in Experiment 1.*

| | Unnecessary Graphics | Relevant Graphics | No Graphics |
|---|---|---|---|
| Pre-Assessment | 0.10 [0.05-0.15] | 0.13 [0.07-0.20] | 0.05 [0.01-0.08] |
| Test | 0.48 [0.43-0.54] | 0.53 [0.47-0.59] | 0.49 [0.44-0.55] |
| Shallow | 0.48 [0.43-0.53] | 0.53 [0.47-0.59] | 0.48 [0.43-0.53] |
| Deep | 0.51 [0.42-0.59] | 0.53 [0.44-0.62] | 0.54 [0.46-0.62] |
| Multiple Choice | 0.48 [0.43-0.53] | 0.54 [0.48-0.60] | 0.49 [0.44-0.54] |
| Open Response | 0.49 [0.42-0.57] | 0.51 [0.43-0.59] | 0.50 [0.42-0.58] |
| Test First Half | 0.47 [0.41-0.54] | 0.47 [0.41-0.54] | 0.49 [0.43-0.55] |
| Test Second Half | 0.50 [0.44-0.56] | 0.55 [0.48-0.62] | 0.50 [0.44-0.56] |
| JOL | 57 [52-62] | 61 [55-66] | 56 [51-62] |

*Mind Wandering.* There was no effect of condition on overall mind wandering, $F(2, 172)$ = 1.34, $p$ = .264, $\eta_p^2$ = .02, $BF_{01}$ = 5.42. There was also no effect of condition for the proportion of reported intentional mind wandering, $F(2, 172) = 0.37$, $p$ = .695, $\eta_p^2$ < .01, $BF_{01}$ = 12.61, but there was a significant difference between our conditions in unintentional mind wandering, $F$(2,

172) = 3.29, $p$ = .039, $\eta_p^2$ = .04, $BF_{10}$ = 0.99, such that those in the no graphics condition reported significantly more unintentional mind wandering than those in both the unnecessary graphics, $t$(118) = 2.27, $p$ = .025, $d$ = 0.41, $BF_{10}$ = 1.94, and the relevant graphics conditions, $t$(112) = 2.19, $p$ = .031, $d$ = 0.41, $BF_{10}$ = 1.68.There was no difference in unintentional mind wandering across the unnecessary and relevant graphics conditions, $t$(114) = 0.13, $p$ = .893, $d$ = 0.02, $BF_{01}$ = 5.02 (see Table 2). That said, the Bayes analysis shows little evidence for this effect.

To investigate time on task effects, mind wandering at the start of the video was defined as mind wandering reported on the first two probes, and mind wandering at the end of the video was defined as mind wandering reported on the last two probes. We conducted a condition by time (start, end) mixed ANOVA on the proportion of overall mind wandering. There was a main effect of time, $F$(1, 172) = 31.80, $p$ < .001, $\eta_p^2$ = .16, $BF_{10}$ = 237,837.74, such that at the end of the video lecture there was a greater proportion of mind wandering, compared to the start. There was no interaction between time and condition, $F$(2, 172) = 0.39, $p$ = .677, $\eta_p^2$ < .01, $BF_{01}$ = 15.79. This pattern of results was the same for both types of mind wandering (intentional: time, $F$(1, 172) = 15.37, $p$ < .001, $\eta_p^2$ = .08, $BF_{10}$ = 168.47, interaction, $F$(2, 172) = 0.79, $p$ = 457, $\eta_p^2$ < .01, $BF_{01}$ = 12.79; unintentional: time, $F$(1, 172) = 13.17, $p$ < .001, $\eta_p^2$ = .07, $BF_{10}$ = 61.97, interaction, $F$(2, 172) = 0.61, $p$ = .546, $\eta_p^2$ < .001, $BF_{01}$ = 10.48; see Table 3).

*Table 2: Means and 95% CIs for the proportion of reported mind wandering during the video lecture for each type of mind wandering by condition in Experiment 1.*

| | Unnecessary Graphics | Relevant Graphics | No Graphics |
|---|---|---|---|
| Overall MW | 0.35 [0.29-0.41] | 0.33 [0.26-0.40] | 0.40 [0.34-0.46] |
| Intentional MW | 0.09 [0.07-0.12] | 0.08 [0.05-0.10] | 0.09 [0.06-0.12] |
| Unintentional MW | 0.24 [0.20-0.28] | 0.24 [0.18-0.29] | 0.31 [0.27-0.36] |

*Distraction Frequency.* We examined the frequency of distraction events which occurred during the video lecture, across the conditions with a one way ANOVA. Three participants were removed for having extreme scores (z scores above 3). There was a significant main effect of condition, $F(2, 169) = 5.19$, $p = .006$, $\eta_p^2 = .06$, $BF_{10} = 4.95$, such that those in the relevant graphics condition had less distraction events during the video ($M = 3.52$), compared to the unnecessary graphics condition ($M = 9.12$), $t(111) = 3.20$, $p = .001$, $d = 0.60$, $BF_{10} = 17.76$, and the no graphics condition ($M = 6.68$), $t(111) = 2.16$, $p = .033$, $d = 0.41$, $BF_{10} = 1.59$. Those in the unnecessary graphics and no graphics condition did not significantly differ, $t(116) = 1.27$, $p = .206$, $d = 0.23$, $BF_{01} = 2.47$ (see Figure 2).

*Helpfulness and Relevance.* Those in the relevant graphics condition ($M = 5.59$) reported that the graphics were more helpful than those in the unnecessary graphics condition ($M$=4.52), $t(113) = 3.33$, $p = .001$, $d = 0.62$, $BF_{10} = 25.48$. Those in the relevant graphics condition, also reported that the graphics were more relevant ($M = 5.70$) compared to those in the unnecessary graphics condition ($M = 4.97$), $t(113) = 2.49$, $p = .014$, $d = 0.47$, $BF_{10} = 3.16$. The Bayes analysis for these both show strong evidence for the effect.

*Table 3: Means and 95% CIs for the proportion of reported mind wandering during the video lecture across time (start, end) and condition in Experiment 1.*

| | Unnecessary Graphics | | Relevant Graphics | | No Graphics | |
|---|---|---|---|---|---|---|
| | Start | End | Start | End | Start | End |
| Overall MW | 0.25 [0.18-0.33] | 0.40 [0.30-0.51] | 0.21 [0.13-0.29] | 0.43 [0.33-0.53] | 0.31 [0.21-0.41] | 0.49 [0.39-0.59] |
| Intentional MW | 0.03 [0.00-0.06] | 0.11 [0.05-0.18] | 0.03 [0.00-0.06] | 0.13 [0.06-0.19] | 0.08 [0.02-0.13] | 0.12 [0.06-0.18] |
| Unintentional MW | 0.22 [0.16-0.29] | 0.29 [0.19-0.38] | 0.18 [0.11-0.26] | 0.30 [0.21-0.39] | 0.23 [0.15-0.30] | 0.37 [0.28-0.47] |

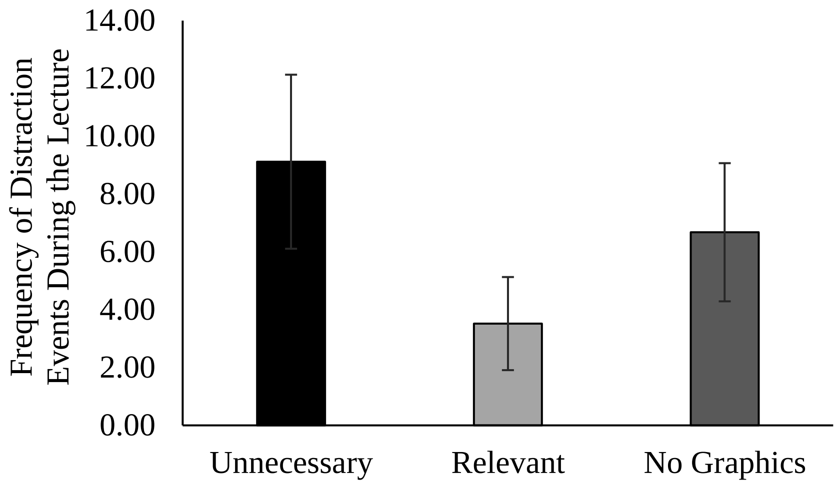


*Figure 2. Average frequency of distraction events during the video lecture by condition (unnecessary, relevant, no graphics), for Experiment 1. Error bars are 95% confidence intervals.*

*2.1.3 Bivariate Correlations*

While not the focus of the present work, we examined the bivariate relations between the various dependent variables, which are displayed in Table 4. Provided the large number of correlations significance values should be interpreted cautiously. We provide the equivalent table of bivariate correlation split by condition in Appendix D. As expected, test accuracy positively correlated with pre-assessment accuracy, and negatively correlated with overall reports of mind wandering and distraction frequency. Test accuracy also positively correlated with JOLs, suggesting that participants could judge their learning, at least to some extent.

Interestingly, mind wandering did not correlate with distraction frequency, but it did negatively correlate with JOLs suggesting that participants may have been aware in some respect that mind wandering relates to reduced learning outcomes. The latter is likely driven by reports of unintentional mind wandering, since JOLs significantly and negatively correlated with unintentional mind wandering but not intentional mind wandering. JOLs also correlated positively with ratings of helpfulness and relevance, these ratings positively correlated with each other. Helpfulness also negatively correlated with overall reports of mind wandering.

*Table 4. Bivariate correlation table for Experiment 1.*

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | |
| 2. Intentional MW | .57** | | | | | | | |
| 3.Unintentional MW | .79** | .20** | | | | | | |
| 4.Test Accuracy | -.18* | .00 | -.12 | | | | | |
| 5. Pre-Assessment Accuracy | -.10 | -.06 | -.07 | .20** | | | | |
| 6. Distraction Frequency | .03 | .02 | .02 | -.22** | -.13 | | | |
| 7. JOL | -.30** | -.12 | -.32** | .37** | .11 | -.00 | | |
| 8. Helpfulness | -.16* | -.13 | -.15 | .03 | .09 | .04 | .17* | |
| 9. Relevance | -.13 | -.12 | -.12 | .06 | .13 | .03 | .18* | .66* |

*Note.* * indicates $p < .05$. ** indicates $p < .01$

*2.1.4 Discussion*

In Experiment 1 there was no effect of the lecture condition on comprehension. In addition, there was no effect of condition on the overall amount of mind wandering, but there was a greater proportion of unintentional mind wandering for those in the no graphics condition compared to conditions with graphics (unnecessary and relevant). JOLs did not significantly vary across the conditions, however participants were overconfident in their estimates. We did find differences across conditions in distraction frequency, such that those in the relevant graphics condition had the lowest rates of being away from the screen compared to the other two lecture conditions. Interestingly, distraction frequency negatively correlated with test accuracy, but did not correlate with mind wandering. Importantly, participants were aware the graphics used in the unnecessary graphics condition were unhelpful and irrelevant to the lecture material, suggesting our conditions were being perceived by participants the way we intended them to be.

Experiment 1 also replicated some results from previous studies. Comprehension and overall mind wandering correlated negatively supporting findings from other studies (Risko et al., 2012). With regards to mind wandering, we found time on task effects for each type of mind wandering, such that there was greater mind wandering reported at the end of the video (Risko et al., 2012). Additionally, participants were overconfident in their JOLs, which is common when JOLs are asked immediately after encoding (Koriat & Bjork, 2006). Before discussing these results further we report a replication and extension.

## 2.2 Experiment 2

### 2.2.1 Rationale

In Experiment 2 we set out to replicate and extend Experiment 1's results. While we did not find differences in comprehension across our lecture conditions in Experiment 1, participant's experiences with the video lecture varied. For example, we found participants were aware the unnecessary graphics were less helpful and relevant than having relevant graphics in the lecture. Some research has suggested that visuals used in a lecture may increase interest, even if it does not improve learning outcomes (Sung & Mayer, 2012; Wilson et al., 2018). For example, Wilson et al. (2018) examined how the addition of an instructor in a video lecture (compared to one with only audio) changed attitudes towards the lecture. They found that participants reported being less likely to drop a class with the instructor in the video, even though this addition did not improve learning outcomes.

While focusing on comprehension is reasonable in an education setting, it is important not to ignore learner's feelings towards the video lectures, since this may influence if they maintain engagement with the material at a coarser level than attending to an individual lecture (Moreno & Mayer, 2007). For example, whether an individual "enjoys" the video lectures in a class would likely influence whether they opened the lecture at all. As such, we included a measure of positive and negative affect for the lecture graphics in Experiment 2.

*2.2.2 Methods*

*Participants.* A total of 226 participants from Amazon's Mechanical Turk participated in this study. Sample size for this study was chosen using the same power analysis as Experiment 1.

*Stimuli.* The video lectures were identical to those used in Experiment 1.

*Measures.* Participants completed the same demographic, pre-assessment, mind wandering probes, JOL and test questions as was used in Experiment 1. In addition to the measures used in Experiment 1, participants completed the positive and negative affect scale (PANAS; Watson, Clark & Tellegen, 1988). The PANAS scale consists of 10 positive word items (e.g., interested) and 10 negative word items (e.g., upset; see Appendix C) that get rated on a 1-5 scale. This scale is reliable for both positive and negative items at alpha >.84. Unlike in the original scale which asks how the participants feel about each of the items in the current moment, we asked participants to "indicate to what extent you felt this way while watching the video lecture".

*Procedure.* The procedure follows Experiment 1 except that immediately after the lecture participants completed the PANAS. They then completed the JOL and the comprehension test. Lastly, they were asked about the helpfulness and relevance of the graphics in the lecture they watched. Once these questions had been completed participants were debriefed and compensated.

*2.2.3 Results*

Thirty four participants were removed for not completing all aspects of the study. An additional 13 participants were removed for failing to correctly answer the question distinguishing the types of mind wandering. The distribution of the remaining 179 participants for each condition were as follows: 61 participants saw unnecessary graphics, 59 saw relevant graphics, and 59 saw no graphics. The mean age was 37 years old. There were 92 males, 81 females, and 1 other. Participants also reported their highest level of education with 63 participants reporting having a high school diploma, 32 having a college diploma, 68 having a bachelor's degree, 13 having a master's degree, and 2 having a doctorate degree. A final question asked was if participants were currently university students, 152 reported not being a current student. The analyses follow those in Experiment 1 with the additional analysis of the positive and negative affect scale.

*Comprehension.* We found no effect of condition for the proportion of correct answers for the pre-assessment questions, $F(2,176) = 0.60$, $p = .551$, $\eta_p^2 < .01$ $BF_{01} = 10.49$, and the majority of participants (82%) had a score of zero on the pre- assessment questions. There was also no effect of condition for the proportion of correct answers to the test questions, $F(2,176) = 2.28$, $p = .106$, $\eta_p^2 = .03$, $BF_{01} = 2.46$ (see Table 5).

Unlike in Experiment 1, when we examined the proportion of correct answers to the shallow test questions there was a significant effect of condition, $F(2, 176) = 3.61$, $p = .029$, $\eta_p^2 = .04$, $BF_{10} = 1.27$, such that those in the unnecessary graphics condition did significantly better on these questions than those in the no graphics condition, $t(118) = 2.68$, $p = .009$, $d = 0.49$, $BF_{10} = 4.67$, but did no better than those in the relevant graphics condition, $t(118) = 1.42$, $p = 0.159$, $d = 0.26$, $BF_{01} = 2.08$. There was no difference between those in the relevant graphics and no graphics conditions $t(116) = 1.25$, $p = .215$, $d = 0.22$, $BF_{01} = 2.54$. That said the Bayes factor is only anecdotally supporting this effect. There was no effect of condition for the deep learning questions, $F(2, 176) = 0.00$, $p = .999$, $\eta_p^2 < .01$, $BF_{01} = 17.61$ (see Table 5).

We also found an effect of condition when we analyze just the multiple choice questions, $F(2, 176) = 3.65$, $p = .028$, $\eta_p^2 = .04$, $BF_{10} = 1.32$, such that those in the unnecessary graphics condition did significantly better on these questions than those in the no graphics condition, $t(118) = 2.66$, $p = .009$, $d = 0.49$, $BF_{01} = 4.52$, but did no better than those in the relevant graphics condition, $t(118) = 1.40$, $p = 0.165$, $d = 0.25$, $BF_{01} = 2.14$. There was no difference between those in the relevant graphics and no graphics conditions, $t(116) = 1.31$, $p = .193$, $d = 0.24$, $BF_{01} = 2.34$. There was no effect of condition for the open response questions, $F(2,176) = 0.34$, $p = .714$, $\eta_p^2 < .01$, $BF_{01} = 13.16$ (see Table 5).

In the analysis including time as a factor, there was a significant main effect of time, $F(1, 176) = 23.98$, $p < .001$, $\eta_p^2 = .12$, $BF_{10} = 6870.52$, such that participants had greater accuracy for questions with content that was from the second half of the video lecture ($M = 0.56$) compared to questions with content from the first half of the lecture ($M = 0.48$). There was no interaction between time and lecture condition, $F(2,176) = 0.02$, $p = .981$, $\eta_p^2 < .01$, $BF_{01} = 17.70$ (see Table 5).

*Table 5: Means and 95% CIs for the proportion of correct answers to the pre-assessment questions, the test questions, the test questions split by type (shallow, deep), test format (multiple choice, short answer) and time (start, end), and participants' JOL (score out of 100) in Experiment 2.*

| | Unnecessary Graphics | Relevant Graphics | No Graphics |
|---|---|---|---|
| Pre-Assessment | 0.10 [0.04-0.15] | 0.06 [0.02-0.10] | 0.08 [0.04-0.13] |
| Test | 0.55 [0.49-0.61] | 0.51 [0.45-0.56] | 0.47 [0.42-0.52] |
| Shallow | 0.57 [0.51-0.63] | 0.51 [0.46-0.57] | 0.47 [0.42-0.57] |
| Deep | 0.48 [0.39-0.56] | 0.47 [0.39-0.56] | 0.48 [0.41-0.55] |
| Multiple Choice | 0.58 [0.52-0.64] | 0.52 [0.47-0.58] | 0.47 [0.43-0.52] |
| Open Response | 0.50 [0.42-0.58] | 0.47 [0.39-0.56] | 0.46 [0.40-0.52] |
| Test First Half | 0.52 [0.46-0.59] | 0.48 [0.41-0.54] | 0.44 [0.38-0.49] |
| Test Second Half | 0.60 [0.54-0.66] | 0.55 [0.49-0.61] | 0.52 [0.45-0.58] |
| JOL | 56 [50-61] | 51 [45-58] | 52 [45-58] |

*Judgements of Learning.* Two participants did not respond to this question and therefore are not included in the analysis. There was no effect of condition for JOLs, $F(2, 174) = 0.65$, $p = .523$, $\eta_p^2 < .01$, $BF_{01} = 9.93$ (see Table 5). There was also no difference between participants' JOLs and their score on the test, $t(176) = 1.02$, $p = .311$, $d = 0.07$, $BF_{01} = 7.18$, thus unlike Experiment 1 participants were not overconfident.

*Mind Wandering.* Consistent with Experiment 1, there was no effect of condition on proportion of overall mind wandering, $F(2, 176) = 0.74$, $p = .478$, $\eta_p^2 < .01$, $BF_{01} = 9.27$, or intentional mind wandering, $F(2, 176) = 0.23$, $p = .791$, $\eta_p^2 < .01$, $BF_{01} = 14.38$,. Unlike in

Experiment 1, we find no effect of condition for unintentional mind wandering, $F(2, 176) = 0.36$, $p = .700$, $\eta_p^2 < .01$, $BF_{01} = 12.94$, (see Table 6).

*Table 6: Means and 95% CIs for the proportion of reported mind wandering during the video lecture for each type of mind wandering by video condition in Experiment 2.*

| | Unnecessary Graphics | Relevant Graphics | No Graphics |
|---|---|---|---|
| Overall MW | 0.41 [0.34-0.47] | 0.45 [0.39-0.52] | 0.46 [0.39-0.52] |
| Intentional MW | 0.12 [0.07-0.16] | 0.15 [0.10-0.17] | 0.13 [0.09-0.18] |
| Unintentional MW | 0.29 [0.24-0.34] | 0.30 [0.25-0.34] | 0.32 [0.27-0.37] |

With respect to time on task, for overall mind wandering there was a main effect of time, $F(1, 176) = 63.27$, $p < .001$, $\eta_p^2 = .26$, $BF_{10} = 5.06e+10$, such that at the end of the video lecture there was more reported mind wandering ($M = 0.58$), compared to the start ($M = 0.32$). There was no interaction between time and condition, $F(2, 176) = 0.47$, $p = .623$, $\eta_p^2 < .01$, $BF_{01} = 13.56$ (see Table 7). This same pattern of results was found for both intentional and unintentional mind wandering types (intentional: time, $F(1, 176) = 26.84$, $p < .001$, $\eta_p^2 = .13$, $BF_{10} = 49{,}783.99$ interaction, $F(2, 176) = 0.44$, $p = .644$, $\eta_p^2 < .01$, $BF_{01} = 11.16$; unintentional: time, $F(1, 176) = 17.82$, $p < .001$, $\eta_p^2 = .09$, $BF_{10} = 515.18$, interaction, $F(2, 176) = 0.19$, $p = .830$, $\eta_p^2 < .01$, $BF_{01} = 15.16$).

*Table 7: Means and 95% CIs for the proportion of reported mind wandering during the video lecture across time (start, end) and condition in Experiment 2.*

| | Unnecessary Graphics | | Relevant Graphics | | No Graphics | |
|---|---|---|---|---|---|---|
| | Start | End | Start | End | Start | End |
| Overall MW | 0.33 [0.24, 0.41] | 0.54 [0.44, 0.64] | 0.28 [0.19, 0.37] | 0.56 [0.45, 0.67] | 0.35 [0.26, 0.44] | 0.63 [0.53, 0.73] |
| Intentional MW | 0.05 [0.01, 0.09] | 0.16 [0.08, 0.23] | 0.07 [0.01, 0.11] | 0.22 [0.13, 0.31] | 0.06 [0.01, 0.11] | 0.18 [0.10, 0.21] |
| Unintentional MW | 0.28 [0.20, 0.36] | 0.39 [0.29, 0.48] | 0.21 [0.13, 0.28] | 0.34 [0.24, 0.43] | 0.29 [0.20, 0.37] | 0.44 [0.34, 0.54] |

*Positive and Negative Affect Scale.* Fourteen participants were removed from this analysis for providing an impossible response (a number that was not between 1 and 5) or for leaving a response on the questionnaire blank. For the remaining 165 participants we calculated their positive and negative affect score and conducted a one-way ANOVA to determine if affect scores varied by condition. There were no differences across the conditions for positive affect ratings, $F(2, 162) = 0.39$, $p = .679$, $\eta_p^2 < .01$, $BF_{01} = 13.80$ (see Table 8) or for negative affect ratings, $F(2, 162) = 1.81$, $p = .166$, $\eta_p^2 = .02$, $BF_{01} = 2.35$ (see Table 8). We examined individual items from the PANAS to determine if ratings for any single item (e.g., interest) varied by condition, however there was no significant effect of condition for any item on the scale (all p's>.1; see Appendix E).

*Table 8: Means and 95% CIs for the Positive and Negative Affect scale by condition in Experiment 2.*

| | Unnecessary Graphics | Relevant Graphics | No Graphics |
|---|---|---|---|
| Positive Affect | 26.77 [24.54-29.00] | 25.70 [23.12-28.28] | 25.36 [22.99-27.72] |
| Negative Affect | 12.43 [11.39-13.46] | 12.08 [10.94-13.21] | 13.82 [12.02-15.62] |

*Distraction Frequency.* Seven participants were removed for having extreme scores (z scores above 3). Unlike in Experiment 1, there was no main effect of condition, $F(2, 169) = 1.36$, $p = .258$, $\eta_p^2 = .02$, $BF_{01} = 5.26$ (see Figure 3).

*Helpfulness and Relevance.* Those in the relevant graphics condition ($M = 5.38$) reported greater helpfulness of the graphics compared to those in the unnecessary graphics condition ($M = 4.75$), $t(116) = 2.20$, $p = .029$, $d = 0.41$, $BF_{10} = 1.71$ as well as greater relevance of the graphics ($M = 5.74$) compared to those in the unnecessary graphics condition ($M = 4.97$), $t(117) = 2.97$, $p$

= .003, *d* = 0.54, $BF_{10}$ = 9.48. As in Experiment 1 the Bayes factors suggest support for this effect, though it is only anecdotal for the difference in helpfulness ratings.

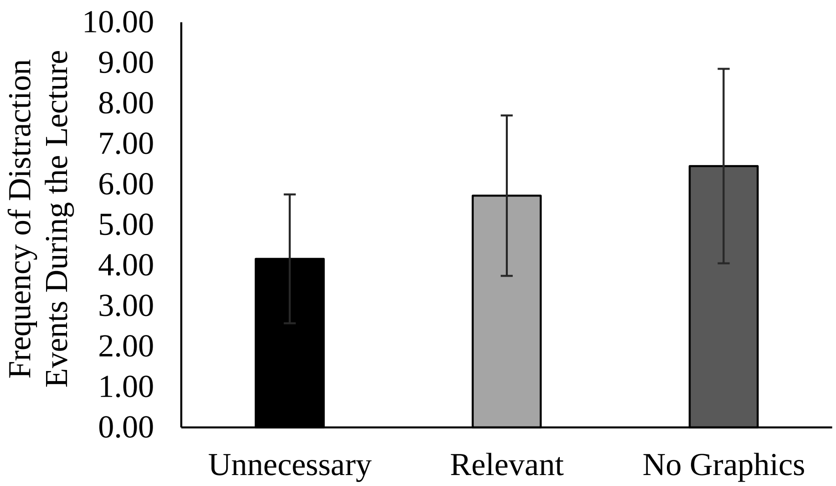


*Figure 3. Average frequency of distraction events during the video lecture by condition (unnecessary, relevant, no graphics), for Experiment 2. Error bars are 95% confidence intervals.*

*2.2.4 Bivariate correlations*

The bivariate correlations between all our dependent variables are displayed in Table 9 Again, correlations should be interpreted cautiously given the number of correlations. The equivalent table of bivariate correlations split by condition are in Appendix D. As with Experiment 1, test accuracy positively correlated with pre-assessment accuracy, and negatively correlated with overall reports of mind wandering and distraction frequency. As well, both types of mind wandering and distraction frequency negatively correlated with JOLs. Again, test accuracy positively correlated with JOLs.

Unlike in Experiment 1, mind wandering did positively correlate with distraction frequency.

Like in Experiment 1, helpfulness ratings negatively correlated with overall mind wandering, however in this Experiment we also find relevance ratings showed the same relation.

Again, both ratings positively correlated with each other suggesting that if the graphics were perceived as helpful they were also perceived to be relevant.

Interestingly, positive, but not negative, affect negatively correlated with both types of mind wandering and distraction frequency, suggesting that greater distraction is associated with less positive affect. Positive and negative affect oppositely related to JOLs, such that positive affect was positively related and negative affect was negatively related. As well, positive affect significantly correlated with participants' pre-assessment scores, suggesting that greater prior knowledge related to greater positive affect. Lastly, positive affect positively correlated with helpfulness and relevance ratings, as well as reports of negative affect.

*Table 9:* Bivariate correlation table for Experiment 2.

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | | | |
| 2. Intentional MW | .60** | | | | | | | | | |
| 3. Unintentional MW | .73** | .00 | | | | | | | | |
| 4. Test Accuracy | -.16* | -.18* | -.08 | | | | | | | |
| 5. Pre-assessment Accuracy | -.10 | -.04 | -.07 | .24** | | | | | | |
| 6. Distraction Frequency | .23** | .22** | .08 | -.16* | -.01 | | | | | |
| 7. JOL | -.40** | -.23** | -.35** | .33** | .20** | -.19** | | | | |
| 8. Helpfulness | -.28** | -.08 | -.24** | .03 | .12 | -.03 | .18* | | | |
| 9, Relevance | -.20** | -.13 | -.11 | .11 | .10 | -.03 | .14 | .68** | | |
| 10. Positive Affect | -.36** | -.21** | -.29** | .06 | .20** | -.24** | .37** | .32** | .20** | |
| 11. Negative Affect | .11 | -.12 | -.00 | -.13 | .07 | .05 | -.20** | -.06 | -.08 | .37** |

*Note.* * indicates $p < .05$. ** indicates $p < .01$.

*2.2.5 Discussion*

As in Experiment 1, in Experiment 2 there were no effect of condition on overall comprehension. That said, those in the unnecessary graphics condition did better on shallow learning questions, and multiple choice questions (these were largely the same items), than did those in the no graphics condition. This was not the case in Experiment 1. Again, there was no effect of condition on the overall amount of mind wandering, but unlike in Experiment 1, there was also no effect of condition on unintentional mind wandering. There was no effect of condition on JOLs and participants were not overconfident in their estimates as we would have expected given what we found in Experiment 1.

In Experiment 2 there was no effect of condition on our exploratory measure of distraction, though we found an effect in Experiment 1. We also found that distraction frequency positively correlated with overall mind wandering, which it did not in Experiment 1. Consistently, distraction frequency and mind wandering negatively correlated with test accuracy. As well, we continue to find evidence of time on task effects for each type of mind wandering, with mind wandering increasing over time.

While participants were aware the graphics used in the unnecessary graphics condition were unhelpful and irrelevant to the lecture material, as they did in Experiment 1, we did not find an effect of condition for ratings of positive and negative affect. That said, positive affect significantly and negatively correlated with both mind wandering and distraction frequency.

## Chapter 3: General Discussion

Across our two experiments we examined the effect unnecessary graphics (i.e., graphics which had no additional explanatory value to the lecture material) had on comprehension and mind wandering in a video lecture compared to a lecture with relevant graphics and one with no graphics. Overall, there were no consistent effects of our video conditions on comprehension of the material overall (nor interactions with time on task). Similarly, for mind wandering, there was no consistent effect of condition on overall or on any specific type of mind wandering (nor interactions with time on task). There was also no effect of condition on participants' JOLs.

When we examined distraction frequency, we did not find a consistent effect of condition. In Experiment 1 the relevant graphics had the lowest frequency of navigating away from the task screen, but this was not replicated in Experiment 2. Similarly, in Experiment 1 there was no correlation between distraction frequency and mind wandering, however there was a positive correlation between them in Experiment 2.

Despite not finding an effect of condition, we did find a number of results that are consistent with existing literature. First, in both experiments there was a negative correlation between overall mind wandering and test accuracy. Additionally, we found that mind wandering occurred often (around 40% of the time), and that reports of mind wandering increased as the time on the task increased.

While there were limited effects of condition on most of our dependant variables, we do find participants ratings of helpfulness and relevance differed reliably across the conditions as we would expect (i.e., the relevant graphics are rated as more helpful and relevant than the unnecessary graphics). While this suggests participants were sensitive to the graphics used, we did not find that the graphics had any effect on ratings of positive and negative affect in Experiment 2.

### *3.1 The Influence of Graphics on Comprehension and Attention*

Overall, there was no effect of condition on comprehension. As noted above, despite there being clear empirical demonstrations of negative effects of irrelevant graphics on comprehension, this effect is likely small and moderated by a number of factors. Thus, it might not be surprising that we found no such effect here. That said, it is possible that our video conditions had no effect due to the fact that our graphics were ineffective. This might also explain why, across our measures of mind wandering, distraction frequency and affect, we found no consistent effect of video condition. That is, the unnecessary graphics might not have been sufficiently distracting or the relevant graphics sufficiently informative (or both). That said, participants were attending to the graphics enough to notice they were unhelpful and irrelevant in the unnecessary condition. It is important to note that in this study, our goal was for ecological validity rather than producing a large effect of graphics, and as such we chose graphics common to what a lecture may actually use (e.g., stock photo of a student reading a book, image of a cat).

Another potential explanation for the lack of effect of video condition on comprehension, could be due to the test questions we asked. Mayer (1999) has suggested that many effects in the multimedia learning literature are specific to “deep” or “transfer” type questions that require learners to engage with the material beyond simple recall. For example, Mayer (1999) summarizes a series of studies that found that “transfer” question performance was affected in lessons with designs that may cause overload (e.g., the use of additional extraneous text or images). While our deep questions required slightly more than verbatim responses from the lecture, they might not have required participants to transfer their knowledge in a new way

(Mayer, 2002). For example, in one of our deeper learning questions we ask what type of reasoning is involved to solve the syllogism, and though the content of this syllogism is different from what was presented in the lecture, it was ordered the same way (i.e., if a then b, a therefore b). Consistent with this idea, in Experiment 1 the deep questions were no harder than the shallow ones, $t(174) = 1.43$, $p = .152$, $d = .11$ (see Table 1), however, participants did significantly worse on deep questions in Experiment 2, $t(178) = 2.41$, $p = .017$, $d = .18$ (see Table 5).

*3.2 Measuring Attention in Lectures*

The present research introduced a novel measure of attentional engagement in online lectures via a behavioral index of whether or not individuals navigate away from the lecture display. Importantly, we consistently found a negative correlation between distraction frequency and test accuracy. This suggests that this behavioural measure of distraction can predict learners' test accuracy. Distraction frequency and mind wandering were not correlated in Experiment 1 but moderately correlated in Experiment 2 suggesting that this measure, at least to some extent, is uniquely measuring a learners' attentional engagement (i.e., distraction frequency may be capturing something beyond what is captured by mind wandering reports). Like mind wandering, distraction frequency was also negatively related to positive affect and JOLs. One interesting possibility is that our distraction measure is capturing learner's media multitasking live, since past research has shown a positive correlation for self-reported media multitasking, and self-reported mind wandering (Ralph et al., 2014).

*3.3 Conclusions*

Across two Experiments, we found no effect of graphics on both overall comprehension and attentional engagement. These results suggest that unnecessary graphics might not be detrimental to comprehension which is consistent with a recent review of this effect, where the authors suggest this negative effect may be exaggerated (Eitel & Kuhl, 2018). From a practical point of view, our results suggest that instructors producing a video lecture may not need to be concerned with the incorporation of some unnecessary graphics. Future work further investigating other relevant instructional design issues from an attention and learning perspective would be valuable.

**References**

Burak, L. (2012). Multitasking in the university classroom. *International Journal for the Scholarship of Teaching and Learning*, *6*, 1-12.

Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, *90*, 1073–1091.

Eitel, A., Bender, L., & Renkl, A. (2019). Are seductive details seductive only when you think they are relevant? An experimental test of the moderating role of perceived relevance. *Applied Cognitive Psychology*, *33*, 20–30.

Eitel, A., & Kühl, T. (2018). Editorial: Harmful or helpful to learning? The impact of seductive details on learning and instruction. *Applied Cognitive Psychology*, 33, 3–8.

Farley, J., Risko, E. F., & Kingstone, A. (2013). Everyday attention and lecture retention: The effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*, 1-9.

Garner, R., Gillingham, M. G., & White, C. S. (1989). Effects of "seductive details" on macroprocessing and microprocessing in adults and children. *Cognition and Instruction*, *6*, 41–57.

Gorissen, P., van Bruggen, J., & Jochems, W. (2012). Students and recorded lectures: Survey on current use and demands for higher education. *Research in Learning Technology*, *20,* 297–311.

Harp, S. F., & Maslich, A. A. (2005). The consequences of including seductive details during lecture. *Teaching of psychology*, *32*, 100-103.

Harp, S. F., & Mayer, R. E. (1997). The role of interest in learning from scientific text and illustrations. On the distinction between emotional interest and cognitive interest. *Journal of Educational Psychology, 89,* 92-102.

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*, 414-434.

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

Ketxer-Noltge, A., Schweppe, J., & Rummer, R. (2018). Is the seductive details effect moderated by mood? An eye tracking study. *Applied Cognitive Psychology, 33*, 62-70.

Koriat, A., & Bjork, R.A. (2006). Illusions of competence during study can be remedied by manipulations that enhance learners' sensitivity to retrieval conditions at test. *Memory and Cognition, 34,* 959-972.

Kühl, T., Moersdorf, F., Römer, M., & Münzer, S. (2018). Adding emotionality to seductive details—Consequences for learning? *Applied Cognitive Psychology, 33*, 48-61.

Magner, U. I. E., Schwonke, R., Aleven, V., Popescu, O., & Renkl, A. (2014). Triggering situational interest by decorative illustrations both fosters and hinders learning in computer-based learning environments. *Learning and Instruction*, *29,* 141–152.

Mayer, R. E. (1999). Multimedia aids to problem-solving transfer. *International Journal of Educational Research, 31*, 611-623.

Mayer, R. E. (2002) Rote versus meaningful learning. *Theory Into Practice, 41*, 226-232.

Mayer, R. E. (2005). Cognitive theory of multimedia learning (1st ed.). *The Cambridge handbook of multimedia learning*, 31–48. New York: Cambridge University Press.

Mayer, R. E. (2009). Multimedia learning (2nd ed.). New York, NY, US: Cambridge University Press.

Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*, 484–490.

Mayer, R. E., Griffith, E., Jurkowitz, I. T. N., & Rothman, D. (2008). Increased interestingness of extraneous details in a multimedia science presentation leads to decreased learning. *Journal of Experimental Psychology: Applied*, *14*, 329–339.

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

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning, *Educational Psychologist*, *38*, 43-52.

McVay, J. C., & Kane, M. J. (2012). Why does working memory capacity predict variation in reading comprehension? On the influence of mind wandering and executive attention. *Journal of Experimental Psychology: General*, *141*, 302-320.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments, contemporary issues and trends. *Educational Psychology Review*, *19*, 309–326.

Moreno, R., & Mayer, R. E. (1999). Cognitive principles of multimedia learning: The role of modality and contiguity. *Journal of Educational Psychology*, *91*, 358-368.

Najjar, L. J. (1996). Multimedia information and learning. *Journal of Educational Multimedia and Hypermedia*, *5*, 129-150.

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

Park, B., Moreno, R., Seufert, T., & Brünken, R. (2011). Does cognitive load moderate the seductive details effect? A multimedia study. *Computers in Human Behavior*, *27,* 5-10.

Ralph, B.C.W., Thomson, D.R., Cheyne, J.A., & Smilek, D. (2014). Media multitasking and failures of attention in everyday life. *Psychological Research, 78,* 661-668.

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*, 234–242.

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

Sanchez, C. A., & Wiley, J. (2006). An examination of the seductive details effect in terms of working memory capacity. *Memory and Cognition, 34*, 344–355.

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

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*, 946-958.

Strobel, B., Grund, S., & Lindner, M. A. (2019). Do seductive details do their damage in the context of graph comprehension? Insights from eye movements. *Applied Cognitive Psychology*, *33*, 95–108.

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

Szpunar, Karl K., Moulton, S. T., & Schacter, D. L. (2013). Mind wandering and education: From the classroom to online learning. *Frontiers in Psychology*, *4*, 1-7.

"Tracking online and distance education in Canadian universities and colleges: 2018", (June 2019). *Canadian Digital Learning Research Association*. Retrieved from: https://onlinelearningsurveycanada.ca/publications-2018

Unsworth, N., McMillan, B.D. (2013). Mind wandering and reading comprehension: Examining the roles of working memory capacity, interest, motivation, and topic interest. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *39*, 832-842.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, *54*, 1063.

Wilson, K. E., Martinez, M., Mills, C., D'Mello, S., Smilek, D., & Risko, E. F. (2018). Instructor presence effect: Liking does not always lead to learning. *Computers and Education*, *122*, 205–220.

**Appendix A** –*Comprehension Test Questions*

List of the 15 test questions in the order participants received them. Questions 1 and 2 were used as the pre-assessment questions. Questions 1, 3, 4, 5, 6, 7, 8, 14 & 15 contained content in the first half of the lecture, the rest of the questions fell in the last 12.5 minutes. The exact timing for the question's content in the video is noted first. "OR" represents the open response questions, "MC" represents the multiple choice questions. Questions with a * next to the type of response were considered deep learning questions.

1) **07:45- OR**. Name 3 different types of problem solving techniques studied by cognitive psychologists.
2) **18:12- OR**. Name 2 different types of reasoning studied by cognitive psychologists.
3) **03:30- MC**. When it comes to reasoning, problem solving, and decision making, researchers have relied on a method that involves the detailed, concurrent, and non-judgmental observation of the contents of your own consciousness as you work on a problem. This method is called:
4) **05:00- MC**. Writing a letter is best considered an example of:
5) **05:40- MC**. Which of the following is NOT one of the reasons cognitive psychologists have focused on well-defined problems?
6) **08:00- OR***. Sally has forgotten her password to log in to her email account. To solve this problem she enters all of her commonly used passwords until she finds the correct one. This problem solving approach is most similar to which one of the approaches discussed in the lecture (please type your response below):
7) **10:20 & 11:20- MC**. A problem space includes an initial state, intermediate state, and goal state, all of which are important parts of these two problem solving techniques:

8) **10:45- MC**. In means ends analysis a permissible move in the problem space is referred to as an:

9) **14:30- MC**. In Gick and Holyoak (1980), participants who were told to use the story of the general to help solve the tumor problem were better able to solve it than participants not told to use the story. This was attributed to the utility of what problem solving technique:

10) **18:45- MC**. The following is an example of what type of reasoning:

    Brian is a university student, Brian lives in a dormitory; therefore all undergrads live in dormitories

11) **19:40- MC**. When individuals solve syllogisms quantifiers like all, none, and some________________

12) **20:50- MC**. Imagine if I gave you a pattern (e.g., 2-5-9) and asked you to generate the rule that was used to generate it. This task would be examining what type of reasoning?

13) **23:30- MC**. The generation of a quasi-pictorial representation of the relationship between the information in the premises and the conclusion when reasoning would be consistent with which approach

14) **04:00- OR***. The lecture discussed both ill-defined and well-defined problems. Provide an example of each that is different from the examples provided in the lecture

15) **18:34- OR***. Aaron argues that all cats are lazy, you have a cat named Boots, and he concludes that Boots is lazy. What kind of reasoning is Aaron using?

**Appendix B**- *Mind Wandering Probe Timings*

The timing of each probe for all participants. The first 6 fell within the first 12.5 minutes of the 25 minute video lecture and the last 4 fell at the end of the video lecture.

Probe 1- 01:04

Probe 2- 03:29

Probe 3- 05:22

Probe 4- 07:26

Probe 5- 10:08

Probe 6- 13:25

Probe 7- 16:17

Probe 8- 18:32

Probe 9- 20:11

Probe 10- 24:01

**Appendix C** *–Positive and Negative Affect Scale (PANAS)*

This scale consists of 20 words that describe different feelings and emotions. Read each item and then enter a number from the scale below next to the word. Indicate to what extent you felt this way while watching the video lecture. Use the following 1-5 scale to record your answers:

1-Very Slightly or Not at All, 2-A Little, 3-Moderately, 4-Quite a Bit, 5-Extremely

1. Interested ______

2. Distressed______

3. Excited______

4. Upset______

5. Strong______

6. Guilty______

7. Scared______

8. Hostile______

9. Enthusiastic______

10. Proud______

11. Irritable______

12. Alert______

13. Ashamed______

14. Inspired______

15. Nervous______

16. Determined______

17. Attentive______

18. Jittery______

19. Active______

20. Afraid______

**Appendix D-** *Magnitude of Correlations*

Below are the bivariate correlations for all of our dependent variables in both Experiments, split across the lecture conditions. The correlations for Experiment 1 are presented first followed by Experiment 2. Correlations with a * beside it indicates $p < .05$, and those with ** indicates $p < .01$.

***Experiment 1: Unnecessary Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | |
| 2. Intentional MW | .61** | | | | | | | |
| 3. Unintentional MW | .75** | .31* | | | | | | |
| 4. Test Accuracy | -.15 | -.02 | -.04 | | | | | |
| 5. Pre- Assessment Accuracy | -.09 | -.11 | -.01 | .24 | | | | |
| 6. Distraction Frequency | -.02 | -.05 | -.04 | -.34** | -.16 | | | |
| 7. JOL | -.17 | -.13 | -.16 | .26* | -.02 | .05 | | |
| 8. Helpfulness | -.21 | -.29* | -.19 | -.17 | .12 | .07 | .06 | |
| 9. Relevance | -.12 | -.24 | -.14 | -.17 | .18 | .07 | -.02 | .71** |

***Experiment 1: Relevant Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | |
| 2. Intentional MW | .47** | | | | | | | |
| 3. Unintentional MW | .77** | .15 | | | | | | |
| 4. Test Accuracy | -.20 | .01 | -.08 | | | | | |
| 5. Pre-Assessment Accuracy | -.08 | -.14 | .02 | .09 | | | | |
| 6. Distraction Frequency | -.13 | -.10 | -.06 | -.09 | -.01 | | | |
| 7. JOL | -.30* | -.04 | -.34* | .45** | .11 | .02 | | |
| 8. Helpfulness | -.14 | .14 | -.24 | .16 | .16 | .17 | .39** | |
| 9. Relevance | -.19 | .01 | -.23 | .19 | .13 | .15 | .47** | .63** |

***Experiment 1: No Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | |
| 2. Intentional MW | .63** | | | | | | | |
| 3. Unintentional MW | .86** | .15 | | | | | | |
| 4. Test Accuracy | -.16 | .03 | -.22 | | | | | |
| 5. Pre-Assessment Accuracy | -.07 | .15 | -.19 | .32* | | | | |
| 6. Distraction Frequency | .32* | .22 | .27* | -.14 | -.25 | | | |
| 7. JOL | -.42** | -.15 | -.44** | .40** | .23 | -.09 | | |
| 8. Helpfulness | -.10 | -.15 | -.03 | .08 | -.13 | .13 | .08 | |
| 9. Relevance | -.05 | -.05 | -.04 | .17 | -.00 | .03 | .17 | .56** |

***Experiment 2: Unnecessary Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | | | |
| 2. Intentional MW | .59** | | | | | | | | | |
| 3. Unintentional MW | .76** | -.08 | | | | | | | | |
| 4. Test Accuracy | -.03 | .01 | -.04 | | | | | | | |
| 5. Pre-Assessment Accuracy | .09 | .10 | .03 | .31* | | | | | | |
| 6. Distraction Frequency | .14 | .13 | .07 | -.11 | .00 | | | | | |
| 7. JOL | -.45** | -.20 | -.40** | .35** | .06 | -.25 | | | | |
| 8. Helpfulness | -.26* | -.17 | -.18 | -.24 | .16 | .12 | .15 | | | |
| 9. Relevance | -.13 | -.18 | -.02 | -.16 | .09 | .07 | .00 | .72** | | |
| 10. Positive Affect | -.34** | -.31* | -.17 | -.02 | .07 | -.20 | .41** | .29* | .23 | |
| 11. Negative Affect | .03 | -.15 | .16 | -.03 | .18 | .16 | -.07 | -.03 | -.04 | .44** |

***Experiment 2: Relevant Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | | | |
| 2. Intentional MW | .57** | | | | | | | | | |
| 3. Unintentional MW | .69** | .17 | | | | | | | | |
| 4. Test Accuracy | -.24 | -.18 | -.27* | | | | | | | |
| 5. Pre-Assessment Accuracy | -.16 | -.02 | -.16 | .34** | | | | | | |
| 6. Distraction Frequency | .17 | .05 | .05 | -.00 | -.15 | | | | | |
| 7. JOL | -.32* | -.35** | -.29* | .40** | .22 | -.04 | | | | |
| 8. Helpfulness | -.31* | .12 | -.31* | .31* | .12 | .00 | .21 | | | |
| 9. Relevance | -.43** | -.14 | -.27* | .33* | .11 | -.02 | .28* | .73** | | |
| 10. Positive Affect | -.30* | -.10 | -.31* | .25 | .29* | -.20 | .32* | .42** | .27* | |
| 11. Negative Affect | .07 | .21 | -.15 | .00 | .05 | -.12 | -.10 | .03 | -.22 | .46** |

***Experiment 2: No Graphics***

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| 1. Overall MW | | | | | | | | | | |
| 2. Intentional MW | .64** | | | | | | | | | |
| 3. Unintentional MW | .73** | -.04 | | | | | | | | |
| 4. Test Accuracy | -.20 | -.40** | .12 | | | | | | | |
| 5. Pre-Assessment Accuracy | -.25 | -.20 | -.14 | .04 | | | | | | |
| 6. Distraction Frequency | .34** | .40** | .09 | -.31* | .06 | | | | | |
| 7. JOL | -.42** | -.18 | -.35** | .21 | .33* | -.26 | | | | |
| 8. Helpfulness | -.34** | -.21 | -.26* | .16 | .10 | -.25 | .26 | | | |
| 9. Relevance | -.16 | -.11 | -.12 | .39** | .17 | -.14 | .24 | .60** | | |
| 10. Positive Affect | -.46** | -.20 | -.43** | -.11 | .27* | -.32* | .38** | .29* | .18 | |
| 11. Negative Affect | .21 | .27* | -.04 | -.30* | -.02 | .06 | -.36** | -.18 | -.03 | .26* |

**Appendix E-** *Statistics for the effect of condition on each PANAS item*

| PANAS Item | One-Way ANOVA results |
|---|---|
| Interested | $F(2, 162) = 2.24, p = .109, \eta_p^2 = .03$ |
| Distressed | $F(2, 162) = 0.39, p = .675, \eta_p^2 < .01$ |
| Excited | $F(2, 162) = 0.26, p = .770, \eta_p^2 < .01$ |
| Upset | $F(2, 162) = 1.93, p = .148, \eta_p^2 = .02$ |
| Strong | $F(2, 162) = 0.06, p = .941, \eta_p^2 < .01$ |
| Guilty | $F(2, 162) = 1.45, p = .237, \eta_p^2 = .02$ |
| Scared | $F(2, 162) = 1.41 , p = .248, \eta_p^2 = .02$ |
| Hostile | $F(2, 162) = 0.59, p = .556, \eta_p^2 < .01$ |
| Enthusiastic | $F(2, 162) = 0.22, p = .802, \eta_p^2 < .01$ |
| Proud | $F(2, 162) = 0.42, p = .655, \eta_p^2 < .01$ |
| Irritable | $F(2, 162) = 1.85 , p = .161, \eta_p^2 = .02$ |
| Alert | $F(2, 162) = 0.90, p = .407, \eta_p^2 = .01$ |
| Ashamed | $F(2, 162) = 0.68, p = .509, \eta_p^2 < .01$ |
| Inspired | $F(2, 162) = 0.84, p = .435, \eta_p^2 = .01$ |
| Nervous | $F(2, 162) = 0.73, p = .486, \eta_p^2 < .01$ |
| Determined | $F(2, 162) = 1.46 , p = .235, \eta_p^2 = .02$ |

| | |
|---|---|
| Attentive | $F(2, 162) = 1.36$ , $p = .252$, $\eta_p^2 = .02$. |
| Jittery | $F(2, 162) = 1.21$, $p = .301$, $\eta_p^2 = .01$ |
| Active | $F(2, 162) = 0.67$, $p = .514$, $\eta_p^2 < .01$ |
| Afraid | $F(2, 162) = 0.89$, $p = .414$, $\eta_p^2 = .01$ |