The Effects of Video Quality on Online Video Lectures

by

Julianna Salvatierra

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2023

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 learning, there is an increasing need to determine the most effective way to present lectures online (e.g., video lectures). While video lectures have several benefits, when delivered online as they typically are, they are vulnerable to video quality issues such as freezing (i.e., disruptions in audio and video playback). Such events have the potential to compromise learning and the lecture experience more broadly. In this study, we manipulated the presence of freezing in video lectures to examine how freezing might impact students' attention, affect, effort, metacognition, and memory for lecture material. Across our four experiments, we found no consistent effects of the freezing condition on learning, load, and metacognition. There were modest impacts of the freezing condition on affect and attention (i.e., mind wandering), particularly with dense bouts of freezing.

Acknowledgements

This work was supported by a SSHRC Partnership Engage Grant with SSIMWAVE. I wish to express my deepest gratitude to Dr. Evan F. Risko for his motivation, immense knowledge, enthusiasm, and patient guidance of this research work. Your enthusiasm for research continues to inspire me.

I would like to thank my readers, Dr. Jonathan Fugelsang and Dr. Daniel Smilek, for providing invaluable insights and thoughtful comments.

I am thankful for the Cognition and Natural Behaviour Laboratory for their constant support and friendship. I deeply appreciate the lab members for supporting me during my "emoting" sessions and "coffee walks". Special thanks to our research assistants who have been so supportive and without whom this work would not be possible.

I would like to thank my family and friends for their endless love and encouragement. Finally, to Reuben, thank you—these words carry all my respect, love, devotion, and gratefulness for you and for all that you do.

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Online learning has been steadily increasing over the last 20 years. For example, even before the COVID-19 pandemic, Canadian post-secondary institutions reported a 10% increase in online course enrolment from the 2015–2016 school year to the 2016–2017 school year (Canadian Digital Learning Research Association, 2019). The closure of academic institutions in response to COVID-19 resulted in a further shift towards online learning. In 2021, 78% of Canadian post-secondary institutions expected a growing trend of fully online learning as a mode of course delivery (Canadian Digital Learning Research Association, 2021). Given the rising use of online learning, there is a need for researchers to investigate current methods of online content delivery and determine ways to optimize student learning and engagement.

In online learning environments, instructors often use recorded lectures as the primary form of content delivery (Brame, 2016; Schacter & Szpunar, 2015). Most recorded lectures can be considered multimedia presentations, which typically involve presenting material in verbal and visual forms (Mayer, 2005). Although there are clear benefits of recorded lectures (e.g., accessibility, flexibility in consumption), these presentations could be vulnerable to video quality issues. In the context of streaming video and audio, a common form of video quality issue is freezing or buffering events. Freezing events are disruptions in audio and video playback. These events are usually caused by an insufficient downloading of data by the server. That is, data download is not fast enough to keep pace with video playback. Insufficient downloading may be due to an unstable internet connection or lack of available bandwidth. In the present investigation, we examined how freezing impacts learning and the broader learning experience in the context of recorded lectures.

1.1 Theoretical Frameworks

We situate the present investigation within two popular theoretical frameworks related to multimedia learning: Mayer's Cognitive Theory of Multimedia Learning (2005) and Moreno's Cognitive-Affective Theory of Learning with Media (2006). According to Mayer's Cognitive Theory of Multimedia Learning (CTML; 2005), during a multimedia presentation, learners receive verbal (e.g., instructor narration) and visual (e.g., lecture slides) input. For effective learning to occur, learners must monitor and integrate this information. This involves holding information in working memory, organizing the information into coherent visual and verbal models, and integrating the information with prior knowledge from long-term memory, resulting in an overall mental model of the to-be-learned material. In the context of this framework, freezing could introduce problems in information integration if it forces learners to hold information in working memory until the presentation resumes, consequently increasing the likelihood of losing such information from working memory. This increase in load might impair the construction of a proper mental model of to-be-learned material. On the other hand, a freezing event could also benefit a learner. For example, during a freezing event, learners could use that period of time to finish processing the learning material (e.g., processing the verbal and visual materials, integrating them, and connecting them with prior knowledge).

While CTML stresses the importance of considering the cognitive processes involved in learning from multimedia, Moreno's Cognitive-Affective Theory of Learning with Media (CATLM; 2006) stresses the importance of both cognitive processes and the affective states involved in learning from multimedia. Specifically, Moreno extends Mayer's (2005) CTML with three additional assumptions: 1) affective mediation—motivational factors impact learning via cognitive engagement; 2) metacognitive mediation—metacognitive factors impact learning via

the regulation of cognitive and affective processes and; 3) individual differences—differences in prior knowledge impact learning. During a multimedia presentation, while learners are selecting, organizing, and integrating information from the lecture, metacognitive and affective factors are impacting learning. For example, positive emotions might increase motivation (Loderer et al., 2018), which in turn increases learning via increased cognitive engagement (e.g., attention). In addition, learners who are able to effectively monitor their cognitive processes (i.e., metacognitive monitoring), can implement effective strategies, and can control their affect and motivation should be better able to learn (i.e., metacognitive control; Bruning et al., 1999). Thus, the CATLM framework draws attention to the need to consider a number of learning relevant variables including affect, metacognition, and cognitive engagement. Indeed, with respect to freezing, it might be these types of variables that are most likely to be directly influenced. For example, research on video analytics and user experience suggest that freezing events can impact user engagement. Dobrian et al. (2011) found that for all types of video content, the buffering ratio (i.e., time spent in buffering/freezing event) had the largest impact on user engagement. There are a number of mechanisms through which this might be observed. For example, the freezing events might be experienced as disruptive or frustrating, leading to negative affect and decreasing cognitive engagement.

1.2 Present Investigation

In the present investigation, we examined how freezing events in an online lecture (i.e., disruptions in video playback of a lecture) impacted a number of learning relevant variables identified in the CTML and CATLM frameworks. We organize these variables here using the acronym LLAMA for learning, load, affect, metacognition, and attention. In the context of our study, we operationalized the measures of these constructs as follows. For learning, we assess

participants' ability to learn from the recorded lecture through a standard content test post-lecture. For load, participants rate how much effort the lecture required and how much effort they chose to invest in the lecture. For affect, participants rate how much they liked the lecture and completed the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988). For metacognition, we assess participants' predictions regarding their performance on the test using aggregate judgments of learning (JOL) that are provided before and after the test. Lastly, for attention, participants were probed during the lecture about whether they were mind wandering and if so, what type (i.e., intentional or unintentional; Risko et al., 2012). 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 et al., 2016). Together this set of variables will provide a comprehensive picture of the impact of freezing on learning from a recorded lecture grounded in existing theory.

Freezing events can vary along multiple dimensions that could, in principle, modulate their impact. We consider two characteristics of freezing events here—their duration (i.e., how long a freezing event occurs) and how they are distributed throughout the lecture. The duration of a freezing event may impact information integration and the overall experience of a lecture. For example, a long disruption in video playback may be worse than a short disruption, especially if the disruption causes learners to hold information in working memory and experience frustration, leading to negative affect and disengagement. The distribution of freezing events may also have similar effects. For example, a concentrated series of freezing events may be worse than a relatively distributed series of freezing events as frequent disruptions could impair the construction of mental models and increase disengagement during a lecture. A

learner's experience can also change throughout a lecture so where the disruptions occur may further modulate their impact. For example, mind wandering increases with time on task (Risko et al., 2012) so a concentrated series of disruptions towards the end of a lecture may increase disengagement.

As such, in the current study, we manipulated the duration of freezing events and how they are distributed throughout the lecture. In particular, across experiments in all freezing conditions, there were a total of 20 freezing events. These freezing events were either 5 seconds or 15 seconds. In addition, these freezing events were distributed either in the first three minutes of the video (front-loaded), throughout the video (uniform), or in the last three minutes of the video (back-loaded). Loading the freezing events toward the beginning or end also created an experience of a higher rate of freezing events (i.e., a dense period of freezing events) in those periods relative to the uniform condition.

Chapter 2: The Effects of Freezing

2.1 Experiment 1

2.1.1 Methods

Participants

The sample size was determined by doubling the approximate number of participants required to detect a small-to-medium effect of freezing on all the dependent variables (Cohen's *d* = .40) with a power of .80 (alpha = .05, two-tailed) in a 2 x 3 mixed analysis of variance (ANOVA). Participants (N = 108) were recruited from the online platform Prolific (reported gender: 50% female, 2% unreported; M_{age} = 35.09 years). All participants completed the study online and received GBP £5.25 as remuneration for the completion of the study. All participants reported their highest level of education (0.93% no formal education, 0.93% primary/elementary schooling, 21% high school diploma, 9% college diploma, 41% bachelor's degree, 21% master's degree, 6% doctorate degree), the number of online courses they have taken (M = 4.64), and their English proficiency in reading (93% high proficiency, 7% moderate proficiency) and listening (94% high proficiency, 6% moderate proficiency).

Stimuli

Each participant was presented with two video lectures. One video lecture was a 10-minute lecture on Ghanes00, a fictional civilization. This lecture contained 13 slides. We created the fictional Ghanes00 lecture to control for previous knowledge and to limit the impact of looking up the answers to the questions on the Internet. The other video lecture was an 8-minute lecture on Music Theory. This lecture contained 7 slides. Both video lectures consisted of a slideshow presentation paired with a video of the instructor explaining the content of the lecture.

There were four different versions of each lecture. One version of the video lectures did not contain any freezing events. The other three versions contained freezing events. In these versions, the video would freeze, and a loading icon would appear in the middle of the screen to denote a freezing event. In each of the three versions, the video had a total of 20 freezing events, each lasting 5 seconds long. These freezing events were distributed either in the first three minutes of the video (front-loaded), throughout the video (uniform), or the last three minutes of the video (back-loaded). In the front-loaded freezing condition, the first freezing event occurred approximately 10 seconds after the video began. After the freezing event occurred, the video would resume for about 10 seconds then the next freezing event would occur. This pattern continued until 20 freezing events had occurred. There would be about 5 or 6 freezing events per minute of the video. In the back-loaded freezing condition, the pattern of events was identical except the first freezing event occurred 10 seconds into the last three minutes of the video. In the uniform freezing condition, the first freezing event occurred approximately 25 seconds after the video began. After the freezing event occurred, the video would resume for about 10 seconds then the next freezing event would occur. This pattern continued until 20 freezing events had occurred. There would be about 2 freezing events per minute of the video.

Measures

Participants completed a demographic questionnaire where they reported their age, gender, highest level of education, number of online courses they had taken, and English proficiency for reading and listening.

During the video lectures, participants received 4 mind wandering probes. The probe appeared at the same time for everyone, however they were presented at different intervals throughout the lecture to prevent participants from anticipating the probe (for probe times, see

Appendix A). Participants were asked if they had been mind wandering in the moments before the probe, and they had 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 each video, participants provided an estimate of their performance on the upcoming assessment (i.e., “How well do you think you will do on the upcoming assessment? That is, what percentage of questions do you think you will get correct?”). They then completed a test on the content of the lecture they just watched. There were 11 test questions. For Ghanes00, there were 7 multiple choice and 4 open response questions. Seven questions were fact retrieval questions and four questions were inference questions. For Music Theory, there were 8 multiple choice and 3 open response questions. Seven questions were also fact retrieval questions and four questions were inference questions (see Appendix B). After the test, participants provided an estimate of their performance. This time, they were asked what percentage of questions they thought they answered correctly.

After the test phase, participants rated each lecture using the following measures: liking, likelihood of watching the lecture in the same format in the future, affect, and effort. For liking, participants rated how much they liked the lecture on a scale from 1–7 (1 = Very much disliked, 7 = Very much liked). For the likelihood of watching the lecture in the same format in the future, participants rated how likely they would be to watch another video lecture that was presented in the same format as the lecture they had just watched on a scale from 1–7 (1 = Very Unlikely, 7 = Very Likely). For affect, we used the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988) to examine their experienced affect during the lecture. The PANAS scale is comprised of two 10-item mood scales: positive affect and negative affect. Using this scale, individuals rate

the extent to which they experience states of pleasant and unpleasant engagement. Positive affect reflects the extent to which a person experiences high energy and pleasurable engagement. It is measured through items such as alert, active, and enthusiastic. In contrast, negative affect reflects the extent to which a person experiences subjective distress and adverse mood. It is measured through items such as anger, nervous, and irritable (Watson et al., 1988). For each of the 10 positive and 10 negative items, participants were asked "how much you felt this way while watching the Ghanes00/Music lecture" on a scale from 1–5 (1 = Very slightly or not at all, 5 = Extremely). For effort, participants rated how much effort each lecture required and how much they chose to invest in each lecture on a scale from 1–9 (1 = Very, very low, 9 = Very, very high).

As a manipulation check, we asked participants at the end of the study if they experienced any video/playback issues when watching the lectures and in which lecture they experienced the issues.

Procedure

After providing consent, participants were asked to complete the demographic questionnaire. Participants were then told that they would be watching two lectures and that their knowledge of lecture material would be tested after each lecture. They were also instructed not to refresh the page as their progress would not be saved.

Participants then proceeded to watch the first video lecture. During the lecture, participants responded to the 4 mind wandering probes located throughout the lecture. After the first video lecture, participants provided their performance estimates before the test. They then completed the test. They then provided the liking, likelihood of watching the lecture in the same format, affect, and effort ratings.

Participants then proceeded to watch the second video lecture, where they followed the same procedure as the first video lecture. At the end of the study, participants were asked if they had searched for the answers to the questions on the Internet, if they had taken notes during the lecture, if there was a reason their data should not be used, and if they had completed the study fully. They were also asked if they experienced any video/playback issues when watching the lectures and in which lecture they experienced the issues. Once these questions had been completed, participants were debriefed and compensated.

2.1.2 Results

This experiment was pre-registered at https://osf.io/chg9v. A total of 146 participants completed the experiment. Thirty-eight participants were excluded from the data analyses based on exclusion criteria (i.e., failed attention checks during the experiment, searched the answers to test questions, took notes during the lecture, did not watch the lectures fully). Eleven participants dropped out of the study. For eight of these participants, the video lecture with freezing events was the first video lecture. For the remaining three participants, the video lecture with freezing events was the second video lecture. Excluded and dropout participants were replaced, thus target sample size was retained. Participants who did not complete the metacognitive ratings were removed from the metacognitive rating analyses (4 participants). Due to random assignment, the number of participants in each condition was not equal. The distribution of the 108 participants according to freezing position condition were as follows: 35 participants were in a front-loaded freezing condition, 36 were in a uniform freezing condition, and 37 were in a back-loaded freezing condition. When asked about video issues, the majority of participants (85%) responded that there were video issues and correctly identified the video that contained freezing events.

The open response test questions were partitioned according to lecture between two coders (i.e., one coder coded the Ghanes00 lecture questions while the other coder coded the Music lecture questions). Participants could receive one mark for a correct response, half a mark for a partially correct response, or zero for an incorrect response. Each question was scored out of 1. All correct multiple-choice responses received a score of 1.

To examine the effect of freezing and freeze position on all variables, we conducted a 2 x 3 mixed ANOVA with freezing (freezing/no freezing) as a within-subjects factor, and freeze position (front-loaded/uniform/back-loaded) as a between-subjects factor. Means and 95% confidence intervals for each dependent variable are shown in Table 1.

Table 1

Means and 95% CIs for dependent variables in Experiment 1.

	Freezing			No Freezing		
	Front-loaded	Uniform	Back-loaded	Front-loaded	Uniform	Back-loaded
Test Performance	58.18 [51.60–64.76]	56.57 [48.68–64.45]	54.91 [49.04–60.79]	57.14 [49.57–64.72]	54.67 [48.26–61.09]	51.35 [45.08–57.62]
Effort Invested	6.23 [5.82–6.64]	6.03 [5.44–6.62]	6.84 [6.51–7.17]	6.43 [5.94–6.92]	6.08 [5.47–6.70]	6.46 [5.94–6.64]
Effort Required	6.86 [6.31–7.40]	7.03 [6.51–7.54]	7.11 [6.52–7.70]	7.17 [6.58–7.76]	6.69 [6.11–7.28]	7.05 [6.34–7.77]
Overall MW	0.34 [0.23–0.46]	0.35 [0.23–0.46]	0.24 [0.15–0.32]	0.26 [0.14–0.38]	0.37 [0.24–0.50]	0.18 [0.09–0.28]
Intentional MW	0.02 [-0.003–0.05]	0.06 [0.02–0.11]	0.007 [-0.01–0.02]	0.03 [-0.01–0.06]	0.09 [0.02–0.16]	0.02 [-0.02–0.06]
Unintentional MW	0.32 [0.21–0.43]	0.28 [0.18–0.39]	0.23 [0.15–0.31]	0.24 [0.13–0.34]	0.28 [0.18–0.38]	0.16 [0.08–0.24]
JOL Before	59.12 [51.95–66.29]	56.76 [48.64–64.89]	64.41 [58.22–70.59]	51.97 [44.02–59.92]	51.76 [42.50–61.02]	54.22 [46.61–61.82]
JOL After	53.63 [44.72–62.55]	51.18 [42.24–60.11]	57.38 [49.18–65.57]	52.70 [43.08–62.32]	53.53 [43.93–63.13]	50.46 [42.32–58.60]
Liking	3.91 [3.42–4.41]	3.69 [3.08–4.31]	3.97 [3.40–4.55]	4.00 [3.37–4.63]	4.31 [3.65–4.96]	3.70 [3.12–4.63]
Watch Again	3.37 [2.78–3.96]	3.06 [2.39–3.72]	3.78 [3.13–4.44]	3.54 [2.82–4.27]	3.47 [2.83–4.11]	3.11 [2.52–3.70]
Positive Affect	2.33 [2.10–2.56]	2.13 [1.83–2.43]	2.31 [2.07–2.55]	2.30 [2.00–2.59]	2.27 [1.94–2.60]	2.18 [1.93–2.44]
Negative Affect	1.30 [1.12–1.47]	1.46 [1.17–1.74]	1.20 [1.11–1.30]	1.29 [1.15–1.43]	1.32 [1.07–1.56]	1.24 [1.11–1.38]

Test Performance

There was no significant effect of freezing, $F(1, 105) = 1.06$, $p = .305$, $\eta^2_G = .003$, freeze position, $F(2, 105) = 0.65$, $p = .522$, $\eta^2_G = .009$, or interaction between freezing and freeze position, $F(2, 105) = 0.13$, $p = .882$, $\eta^2_G = .001$, on the proportion of correct answers to the content test questions.

Effort

There was no significant effect of freezing, $F(1, 105) = 0.09$, $p = .772$, η^2_G □ .001, freeze position, $F(2, 105) = 1.94$, $p = .149$, $\eta^2_G = .027$, or interaction between freezing and freeze position, $F(2, 105) = 1.54$, $p = .219$, $\eta^2_G = .007$, on invested effort. This pattern of results was the same for required effort (main effects: freezing, $F(1, 105) = 0.02$, $p = .897$, η^2_G □ .001, freeze position, $F(2, 105) = 0.21$, $p = .808$, $\eta^2_G = .003$; interaction, $F(2, 105) = 0.99$, $p = .376$, $\eta^2_G = .006$).

Mind Wandering

There was no significant effect of freezing, $F(1, 105) = 1.65$, $p = .202$, $\eta^2_G = .003$, freeze position, $F(2, 105) = 2.43$, $p = .093$, $\eta^2_G = .035$, or interaction between freezing and freeze position on overall mind wandering, $F(2, 105) = 1.06$, $p = .352$, $\eta^2_G = .004$. This pattern of results was the same for intentional mind wandering (main effects: freezing, $F(1, 105) = 1.06$, $p = .305$, $\eta^2_G = .004$; interaction, $F(2, 105) = 0.15$, $p = .860$, $\eta^2_G = .001$) with the exception of freeze position, $F(2, 105) = 4.42$, $p = .014$, $\eta^2_G = .047$. Participants in the uniform condition ($M = 0.08$) reported more intentional mind wandering compared to the front-loaded ($M = 0.03$), $t(104) = 2.21$, $p = .029$, $d = 0.37$, and back-loaded ($M = 0.01$) conditions, $t(106) = 2.69$, $p = .008$, $d = 0.45$. There was no difference between those in the front-loaded and back-loaded conditions,

$t(141) = 0.77$, $p = .441$, $d = 0.13$. There was a significant effect of freezing on unintentional mind wandering, $F(1, 105) = 4.70$, $p = .032$, $\eta_G^2 = .008$, such that participants reported more mind wandering in the freezing condition compared to the no freezing condition. There was no significant effect of freeze position, $F(2, 105) = 1.24$, $p = .294$, $\eta_G^2 = .019$, and no interaction between freezing and freeze position on unintentional mind wandering, $F(2, 105) = 0.93$, $p = .399$, $\eta_G^2 = .003$.

In creating the front-loaded and back-loaded conditions, we also created time periods of dense freezing episodes. That is, in the front- and back-loaded conditions, all of the freezing events occurred in the first or last 3 minutes respectively, approximately 5 or 6 freezing events per minute. In the uniform condition, on the other hand, the same number of freezing events were distributed across the entire lecture and, as a result, only occurred 2 per minute. In addition, one of our probes in each lecture was placed at a time such that it would occur during these dense periods of freezing thus allowing us to examine how such frequent episodes impacted mind wandering. To examine this, we created two new variables from the front-loaded and back-loaded conditions. The "dense freezing" condition included the mind wandering probes that appeared in areas where freezing events were highly concentrated (i.e., early probes in the front-loaded condition and late probes in the back-loaded condition). The "no freezing" condition included the probes that appeared in areas where freezing events did not occur (i.e., late probes in the front-loaded condition and early probes in the back-loaded condition). Note that the probe locations were identical across these conditions. That is, in the front-loaded condition, the early probe would be considered as "dense", and the late probe would be considered as "no freezing". In the back-load condition, the early probe would be considered "no freezing", and the late probe would be considered as "dense." To examine the effect of freeze density on mind wandering, we

conducted an independent samples t-test comparing the two conditions (dense freezing vs. no freezing). Means and 95% CIs for the proportion of reported mind wandering are shown in Table 2. For overall mind wandering, there was no significant effect of freeze density, $t(141) = 1.40$, $p = .163$, $d = 0.23$. This pattern of results was the same for intentional mind wandering, $t(128) = 0.58$, $p = .563$, $d = 0.10$, and unintentional mind wandering, $t(141) = 1.24$, $p = .216$, $d = 0.21$.

Table 2

Means and 95% CIs for the proportion of reported mind wandering during the video lecture across freeze density (dense freezing, no freezing) in Experiment 1.

	Dense Freezing	No Freezing
Overall MW	0.40 [0.29–0.52]	0.29 [0.18–0.40]
Intentional MW	0.03 [-0.01–0.07]	0.01 [-0.01–0.04]
Unintentional MW	0.38 [0.26–0.49]	0.28 [0.17–0.38]

Metacognition

There was a significant effect of freezing on predicted scores before test (i.e., JOL Before), $F(1, 101) = 7.28$, $p = .008$, $\eta_G^2 = .028$, such that participants reported a higher predicted score in the freezing condition compared to the no freezing condition. There was no significant effect of freeze position, $F(2, 101) = 0.81$, $p = .449$, $\eta_G^2 = .010$, and no interaction between freezing and freeze position, $F(2, 101) = 0.31$, $p = .737$, $\eta_G^2 = .002$, on predicted scores before test. There was no significant effect of freezing, $F(1, 101) = 0.53$, $p = .467$, $\eta_G^2 = .001$, freeze position, $F(2, 101) = 0.04$, $p = .957$, $\eta_G^2 = .001$, or interaction between freezing and freeze position, $F(2, 101) = 1.20$, $p = .305$, $\eta_G^2 = .006$, on predicted scores after test (i.e., JOL After).

We examined the relationship between predicted scores and performance by calculating a Pearson correlation coefficient between prospective/retrospective predicted scores and actual test performance by freezing condition. For predicted scores before test, there was a positive correlation between prospective predicted scores and actual test performance in the freezing, $r(102) = 0.36$, $p < .001$, and no freezing condition, $r(102) = 0.45$, $p < .001$. For predicted scores after test, there was a positive correlation between retrospective predicted scores and actual test performance in the freezing condition, $r(102) = 0.56$, $p < .001$, and the no freezing condition, $r(102) = 0.61$, $p < .001$.

Affect

Liking. There was no significant effect of freezing, $F(1, 105) = 0.47$, $p = .495$, $\eta_G^2 = .002$, freeze position, $F(2, 105) = 0.13$, $p = .875$, $\eta_G^2 = .002$, or interaction between freezing and freeze position, $F(2, 105) = 1.54$, $p = .219$, $\eta_G^2 = .011$, on liking.

Likelihood of watching the lecture in the same format. There was no significant effect of freezing, $F(1, 105) = 0.02$, $p = .892$, η_G^2 □ .001, freeze position, $F(2, 105) = 0.18$, $p = .838$, $\eta_G^2 = .002$, or interaction between freezing and freeze position, $F(2, 105) = 2.40$, $p = .096$, $\eta_G^2 = .015$, on the likelihood of watching the lecture in the same format.

PANAS. There was no significant effect of freezing, $F(1, 105) = 0.01$, $p = .926$, η_G^2 □ .001, freeze position, $F(2, 105) = 0.24$, $p = .789$, $\eta_G^2 = .003$, or interaction between freezing and freeze position, $F(2, 105) = 1.00$, $p = .371$, $\eta_G^2 = .005$, on positive affect. This pattern of results was the same for negative affect (main effects: freezing, $F(1, 105) = 0.69$, $p = .409$, $\eta_G^2 = .001$, freeze position, $F(2, 105) = 0.91$, $p = .407$, $\eta_G^2 = .014$; interaction, $F(2, 105) = 1.57$, $p = .214$, $\eta_G^2 = .005$).

Order Effects

Provided freezing was manipulated within-participants, we included order as a factor in a series of ANOVAs mirroring those reported above. There were a number of interactions with order, but they did not qualitatively alter the conclusions.

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 3. Provided the large number of correlations, significance values should be interpreted cautiously. That said, the correlations were generally consistent with both the CTML and CATLM frameworks (e.g., test performance was positively correlated with invested effort and with liking and affect, and negatively related to required effort, negative affect and mind wandering).

Table 3

Bivariate correlation table for Experiment 1

Variable	1	2	3	4	5	6	7	8	9	10	11
1. Test Performance											
2. Effort Invested	.24**										
3. Effort Required	-.23**	.25**									
4. Overall MW	-.19**	-.53**	.01								
5. Intentional MW	-.10	-.37**	.04	.47**							
6. Unintentional MW	-.17*	-.43**	.01	.92**	.09						
7. JOL Before	.32**	.19**	-.24**	-.22**	-.03	-.23**					
8. JOL After	.49**	.24**	-.28**	-.27**	-.09	-.26**	.79**				
9. Liking	.34**	.33**	-.27**	-.37**	-.23**	-.31**	.29**	.36**			
10. Watch Again	.28**	.33**	-.18**	-.36**	-.19**	-.32**	.30**	.32**	.78**		
11. Positive Affect	.23**	.50**	-.001	-.46**	-.23**	-.41**	.22**	.25**	.67**	.65**	
12. Negative Affect	-.16*	-.17*	.24**	.17*	.12*	.14*	-.15*	-.21**	-.40**	-.31**	-.14*

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

2.1.3 Discussion

In Experiment 1, there were no significant effects of freezing or the freeze position conditions on test performance, effort, or affect. However, freezing did influence mind wandering. While there was no overall effect of freezing on overall mind wandering and intentional mind wandering, there was a significant effect on unintentional mind wandering. Participants reported more unintentional mind wandering in the freezing condition compared to the no freezing condition. With respect to the effect of freeze density, there was no significant effect on mind wandering. However, there seemed to be a trend where participants reported more overall mind wandering at points where dense freezing occurred compared to the points where no freezing occurred. Lastly, there was an effect of freezing in predicted scores before test but not after test. Participants reported higher predicted scores before test in the freezing condition compared to the no freezing condition.

2.2 Experiment 2

In Experiment 2, we sought to replicate the results of Experiment 1 in a between-subjects design. While the order effects did not qualitatively alter the conclusions, we wanted a clearer picture of the effects of freezing or lack thereof. We also revised the liking question to include ratings of four specific aspects of the video lecture: general experience, content, visuals, and instructor.

2.2.1 Methods

Participants

An a-priori sample size was determined to detect a medium effect of freezing on all the dependent variables (Cohen's d = .42) with a power of .80 (alpha = .05, two-tailed) in an

independent-samples *t*-test. Participants (N = 180) were recruited from the online platform Prolific (reported gender: 53% female, 1% unreported; M_{age} = 37.5 years). All participants completed the study online and received GBP £3.00 as remuneration for the completion of the study. All participants reported their highest level of education (1% no formal education, 1% primary/elementary schooling, 21% high school diploma, 11% college diploma, 38% bachelor's degree, 22% master's degree, 6% doctorate degree), the number of online courses they have taken (M = 4.53), and their English proficiency in reading (93% high proficiency, 7% moderate proficiency) and listening (91% high proficiency, 8% moderate proficiency, 1% low proficiency).

Stimuli

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

Measures

Participants completed the same demographic questionnaire, mind wandering probes, content test questions and rating measures (i.e., likelihood of watching the lecture in the same format, affect, and effort) that were used in Experiment 1 except for liking and the manipulation check questions. On a scale from 1–7 (1 = Very much disliked, 7 = Very much liked), participants rated how much they liked the general experience of watching the lecture, the content of the lecture, the visuals used in the lecture, and the instructor in the lecture. As a manipulation check, we only asked participants if they experienced any video/playback issues when watching the lectures.

Procedure

The procedure was similar to Experiment 1 with two exceptions. First, participants were told that they would be watching only one lecture. This lecture either contained freezing events or no freezing events. Second, after the lecture, participants completed the rating measures before the performance estimates and the subsequent test.

2.2.2 Results

This experiment was pre-registered at https://osf.io/zefnq. A total of 200 participants completed the experiment. Twenty participants were excluded from the data analyses based on exclusion criteria (i.e., failed attention checks during the experiment, searched the answers to test questions, took notes during the lecture, did not watch the lectures fully). Eleven participants dropped out of the study. Five of the participants were in the freezing condition while six of the participants were in the no freezing condition. For the five participants in the freezing condition, one participant was in a front-loaded freezing condition, two participants were in a uniform freezing condition, and two participants were in a back-loaded freezing condition. Excluded and dropout participants were replaced, thus target sample size was retained. The distribution of the 180 participants for each condition were as follows: 90 participants were in a freezing condition and 90 were in a no freezing condition. Of the 90 participants in the freezing condition, 31 saw a front-loaded freezing condition, 30 saw a uniform freezing condition, and 29 saw a back-loaded freezing condition. When asked about video quality issues, the majority of participants in the freezing condition (87%) responded that the video did contain freezing events. The majority of participants in the no freezing condition (97%) responded that the video did not contain freezing events.

The coding process for open response test questions was identical to Experiment 1. To examine the effect of freezing on all variables, we conducted an independent-samples *t*-test

comparing the two freezing conditions (freezing/no freezing). To examine the effect of freeze position on all variables, we conducted a one-way ANOVA comparing the three freezing position conditions (front-loaded/uniform/back-loaded). Means and 95% confidence intervals for each dependent variable are shown by freezing condition in Table 4, and by freeze position condition in Table 5.

Table 4

Means and 95% CIs for dependent variables by freezing condition in Experiment 2.

	Freezing	No Freezing
Test Performance	51.11 [46.57–55.66]	53.38 [49.25–57.52]
Effort Invested	6.42 [6.10–6.75]	6.46 [6.18–6.73]
Effort Required	6.39 [6.01–6.77]	6.79 [6.43–7.15]
Overall MW	0.34 [0.28–0.40]	0.30 [0.23–0.37]
Intentional MW	0.04 [0.01–0.06]	0.03 [0.01–0.06]
Unintentional MW	0.30 [0.24–0.36]	0.27 [0.20–0.33]
JOL Before	56.31 [51.50–61.12]	56.16 [51.58–60.73]
JOL After	50.16 [44.58–55.73]	53.13 [47.92–58.35]
Liking (Experience)	4.09 [3.71–4.47]	3.53 [3.17–3.90]
Liking (Content)	4.53 [4.17–4.90]	3.92 [3.52–4.32]
Liking (Visuals)	3.14 [2.82–3.47]	2.76 [2.45–3.06]
Liking (Instructor)	4.84 [4.55–5.14]	4.49 [4.19–4.79]
Watch Again	3.48 [3.06–3.89]	2.82 [2.46–3.19]
Positive Affect	2.31 [2.14–2.48]	2.19 [2.03–2.34]
Negative Affect	1.31 [1.21–1.41]	1.20 [1.13–1.26]

Table 5

Means and 95% CIs for dependent variables by freeze position condition in Experiment 2.

	Freezing		
	Front-loaded	Uniform	Back-loaded
Test Performance	51.76 [44.27–59.25]	56.06 [47.73–64.39]	45.30 [37.03–53.57]
Effort Invested	6.61 [6.11–7.12]	6.37 [5.77–6.97]	6.28 [5.64–6.91]
Effort Required	6.48 [5.96–7.01]	6.33 [5.61–7.06]	6.34 [5.56–7.13]
Overall MW	0.29 [0.20–0.39]	0.31 [0.19–0.42]	0.42 [0.30–0.54]
Intentional MW	0.02 [-0.007–0.04]	0.06 [-0.005–0.12]	0.04 [-0.001–0.09]
Unintentional MW	0.27 [0.18–0.36]	0.25 [0.14–0.36]	0.38 [0.26–0.49]
JOL Before	58.55 [50.74–66.36]	58.67 [50.09–67.24]	51.48 [42.09–60.88]
JOL After	51.52 [42.03–61.00]	53.33 [44.04–62.63]	45.41 [34.26–56.57]
Liking (Experience)	4.52 [3.93–5.11]	4.00 [3.29–4.71]	3.72 [3.02–4.43]
Liking (Content)	4.97 [4.45–5.48]	4.30 [3.59–5.01]	4.31 [3.61–5.01]
Liking (Visuals)	3.68 [3.17–4.18]	2.83 [2.20–3.46]	2.90 [2.32–3.47]
Liking (Instructor)	5.13 [4.71–5.55]	4.70 [4.17–5.23]	4.69 [4.09–5.29]
Watch Again	4.32 [3.60–5.04]	3.03 [2.34–3.72]	3.03 [2.31–3.76]
Positive Affect	2.43 [2.10–2.76]	2.31 [1.98–2.63]	2.18 [1.94–2.43]
Negative Affect	1.20 [1.09–1.32]	1.33 [1.15–1.52]	1.39 [1.17–1.62]

Test Performance

There was no significant effect of freezing $t(176) = 0.73$, $p = .464$, $d = 0.11$, or freeze position, $F(2, 87) = 1.87$, $p = .160$, $\eta_G^2 = .041$, on the proportion of correct answers to the content test questions.

Effort

There was no significant effect of freezing $t(174) = 0.16$, $p = .877$, $d = -0.02$, or freeze position, $F(2, 87) = 0.38$, $p = .684$, $\eta_G^2 = .009$, on invested effort. This pattern of results was the same for required effort (freezing: $t(177) = 1.53$, $p = .129$, $d = -0.23$; freeze position: $F(2, 87) = 0.06$, $p = .938$, $\eta_G^2 = .001$).

Mind Wandering

There was no significant effect of freezing, $t(176) = 0.84$, $p = .402$, $d = 0.13$, or freeze position, $F(2, 87) = 1.73$, $p = .183$, $\eta_G^2 = .038$, on overall mind wandering. This pattern of results was the same for intentional mind wandering (freezing: $t(177) = 0.31$, $p = .760$, $d = 0.05$; freeze position: $F(2, 87) = 0.91$, $p = .407$, $\eta_G^2 = .020$) and unintentional mind wandering (freezing: $t(177) = 0.78$, $p = .438$, $d = 0.12$; freeze position: $F(2, 87) = 1.79$, $p = .173$, $\eta_G^2 = .040$).

Similar to Experiment 1, we investigated the effect of freeze density on mind wandering. Means and 95% CIs for the proportion of reported mind wandering are shown in Table 6. For overall mind wandering, there was no significant effect of freeze density, $t(118) = 0.18$, $p = .854$, $d = 0.03$. This pattern of results was the same for intentional mind wandering, $t(118) = 0.00$, $p = 1$, $d = 0$, and unintentional mind wandering, $t(117) = 0.19$, $p = .852$, $d = 0.03$.

Table 6

Means and 95% CIs for the proportion of reported mind wandering during the video lecture across freeze density (dense freezing, no freezing) in Experiment 2.

	Dense Freezing	No Freezing
Overall MW	0.40 [0.27–0.53]	0.42 [0.29–0.55]
Intentional MW	0.03 [-0.01–0.08]	0.03 [-0.01–0.08]
Unintentional MW	0.37 [0.24–0.49]	0.28 [0.26–0.51]

Metacognition

There was no significant effect of freezing, $t(177) = 0.05$, $p = .963$, $d = 0.007$, or freeze position, $F(2, 87) = 0.94$, $p = .393$, $\eta^2_G = .021$, on predicted scores before test. This pattern of results was the same for predicted scores after test (freezing: $t(177) = 0.77$, $p = .440$, $d = -0.12$; freeze position: $F(2, 87) = 0.71$, $p = .495$, $\eta^2_G = .016$).

For predicted scores before test, there was a positive correlation between prospective predicted scores and actual test performance in the freezing, $r(88) = 0.55$, $p < .001$, and no freezing condition, $r(88) = 0.48$, $p < .001$.

For predicted scores after test, there was a positive correlation between retrospective predicted scores and actual test performance in the freezing, $r(88) = 0.74$, $p < .001$, and no freezing condition , $r(88) = 0.67$, $p < .001$.

Affect

Liking. There was a significant effect of freezing on liking the general experience of the lecture, $t(177) = 2.09$, $p = .038$, $d = 0.31$, such that participants in the freezing condition reported

higher liking for the general experience of lecture compared to participants in the no freezing condition. There was no significant effect of freeze position on liking the general experience of the lecture, $F(2, 87) = 1.52$, $p = .225$, $\eta^2_G = .034$. This pattern of results was the same for liking the content of the lecture (freezing: $t(176) = 2.26$, $p = .025$, $d = 0.34$; freeze position: $F(2, 87) = 1.51$, $p = .227$, $\eta^2_G = .034$). Participants in the freezing condition reported higher liking for the content of the lecture compared to participants in the no freezing condition.

There were no significant effects of freezing or freeze position on liking the visuals of the lecture (freezing: $t(177) = 1.72$, $p = .087$, $d = 0.26$; freeze position: $F(2, 87) = 2.88$, $p = .061$, $\eta^2_G = .062$), and on liking the instructor in the lecture (freezing: $t(177) = 1.70$, $p = .092$, $d = 0.25$; freeze position: $F(2, 87) = 0.99$, $p = .375$, $\eta^2_G = .022$).

Likelihood of watching the lecture in the same format. There was a significant effect of freezing on the likelihood of watching the lecture in the same format, $t(175) = 2.35$, $p = .020$, $d = 0.35$, such that participants in the freezing condition were more likely to watch the lecture in the same format compared to participants in the no freezing condition. There was also a significant effect of freeze position, $F(2, 87) = 4.61$, $p = .013$, $\eta^2_G = .096$. Participants in the front-loaded condition ($M = 4.3$) were more likely to watch the lecture in the same format compared to participants in the uniform ($M = 3.0$), $t(58) = 2.64$, $p = .011$, $d = 0.67$, and back-loaded conditions ($M = 3.0$), $t(57) = 2.56$, $p = .013$, $d = 0.66$. There was no difference between those in the uniform and back-loaded conditions, $t(56) = 0.002$, $p = .998$, $d < 0.01$.

PANAS. There were no significant effects of freezing or freeze position on positive affect (freezing: $t(175) = 1.08$, $p = .280$, $d = 0.16$; freeze position: $F(2, 87) = 0.70$, $p = .501$, $\eta^2_G = .016$), or negative affect (freezing: $t(152) = 1.81$, $p = .073$, $d = 0.27$; freeze position: $F(2, 87) = 1.22$, $p = .301$, $\eta^2_G = .027$).

Bivariate Correlations

We examined the bivariate relations between the various dependent variables, which are displayed in Table 7. Again, significance values should be interpreted cautiously given the number of correlations. As with Experiment 1, the correlations were still generally consistent with both the CTML and CATLM frameworks (e.g., test performance was positively correlated with liking and affect, and negatively related to required effort, negative affect and mind wandering).

Table 7

Bivariate correlation table for Experiment 2

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Test Performance														
2. Effort Invested	.11													
3. Effort Required	-.20**	.38**												
4. Overall MW	-.24**	-.53**	-.01											
5. Intentional MW	-.29**	-.32**	-.07	.38**										
6. Unintentional MW	-.14*	-.44**	.01	.92**	-.015									
7. JOL Before	.52**	.22**	-.32**	-.33**	-.17*	-.28**								
8. JOL After	.71**	.16*	-.26**	-.35**	-.27**	-.27**	.79**							
9. Liking (Experience)	.43**	.29**	-.21**	-.42**	-.21**	-.36**	.69**	.52**						
10. Liking (Content)	.46**	.27**	.02*	-.39**	-.22**	-.33**	.64**	.57**	.83**					
11. Liking (Visuals)	.22**	.19*	-.18*	-.32**	-.14	-.29**	.25**	.27**	.54**	.39**				
12. Liking (Instructor)	.24**	.21**	-.07	-.26**	.001	-.27**	.31**	.26**	.57**	.46**	.52**			
13. Watch Again	.26**	.27**	-.21**	-.42**	-.12	-.40**	.36**	.34**	.72**	.62**	.72**	.56**		
14. Positive Affect	.21**	.64**	.05	-.55**	-.17*	-.52**	.41**	.34**	.57**	.58**	.44**	.46**	.57**	
15. Negative Affect	-.23**	.07	.18*	.04	.00	.04	-.33**	-.33**	-.34**	-.26**	-.27**	-.29**	-.28**	-.05

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

2.2.3 Discussion

As with Experiment 1, in Experiment 2 there were no significant effects of freezing or the freeze position conditions on test performance, positive and negative affect, and effort. Unlike Experiment 1, freezing did not influence mind wandering in this experiment. With respect to the effect of freeze density, there was no significant effect on mind wandering. There were also no effects of freezing on predicted scores before and after test but the correlations between predicted scores and actual test performance were positive. There were no effects of freezing on liking the visuals and instructor, but there were significant effects on liking the general experience and content of the lecture as well as the likelihood of watching the lecture in the same format, such that participants in the freezing condition reported higher ratings for these variables compared to participants in the no freezing condition. There was also a significant effect of freeze position on the likelihood of watching the lecture in the same format, such that participants in the front-loaded condition were more likely to watch the lecture in the same format compared to participants in the uniform and back-loaded conditions.

2.3 Experiment 3

Overall, there were no consistent effects of freezing in Experiments 1 and 2 and this was, in general, not moderated by freezing position. In Experiment 3, we used the same design as in Experiment 2 but increased the duration of the freezing events from 5 seconds to 15 seconds. In addition, one consideration when thinking about the impact of freezing in the context of video lectures is that freezing typically consists of a pause in the delivery of new content. In Experiment 2, participants in the freezing condition still had access to visual information (i.e., lecture slides) that may have helped mitigate the potential costs or amplify the potential benefits due to freezing. As such, in Experiment 3, we also wanted to examine if the impact of freezing

varies as a function of available visual content (i.e., lecture slides). Thus, we manipulated the presence of slides in the freezing and no freezing conditions.

2.3.1 Methods

Participants

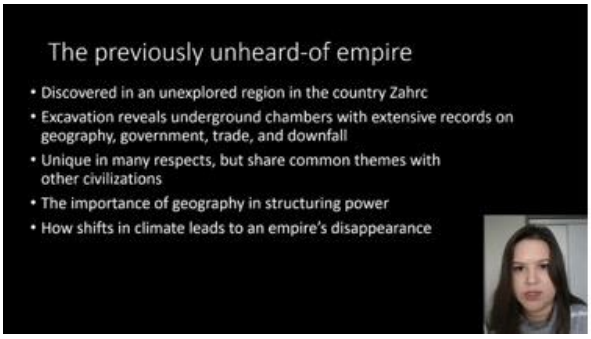
An a-priori sample size was determined to detect a medium effect of freezing on all the dependent variables (Cohen's d = .42) with a power of .80 (alpha = .05, two-tailed) in a 2 x 2 between-subjects ANOVA. Participants (N = 360) were recruited from the online platform Prolific (reported gender: 52% female, 1% unreported; M_{age} = 38.17 years). All participants completed the study online and received GBP £3.45 as remuneration for the completion of the study. All participants reported their highest level of education (0.28% no formal education, 0.28% primary/elementary schooling, 22% high school diploma, 17% college diploma, 41% bachelor's degree, 15% master's degree, 4% doctorate degree), the number of online courses they have taken (M = 5.59), and their English proficiency in reading (94% high proficiency, 5% moderate proficiency) and listening (93% high proficiency, 7% moderate proficiency).

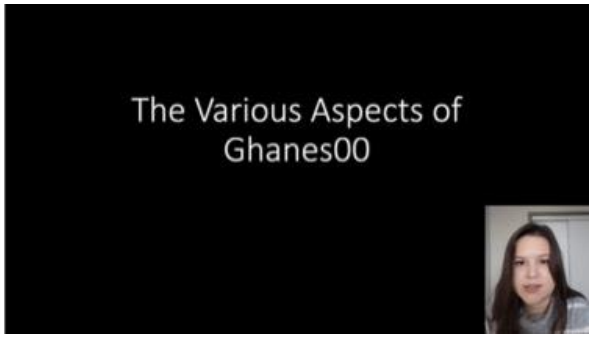
Stimuli

The video lectures were similar to those used in Experiment 1 and Experiment 2 with two exceptions. First, the duration of the freezing events increased from 5 seconds to 15 seconds. The amount and distribution of the freezing events were the same. Second, the video lectures either contained visual content (i.e., lecture slides with information) or did not contain visual content (i.e., lecture slides with only the lecture title).

Figure 1

Depiction of stimuli in Experiment 3





Note. In the slides condition (left), the video lecture would contain lecture slides with information. In the no slides condition (right), the video lecture would contain lecture slides with only the lecture title.

Measures

Participants completed the same demographic questionnaire, mind wandering probes, content test questions and rating measures (i.e., liking, likelihood of watching the lecture in the same format, affect, and effort) that were used in Experiment 2.

Procedure

The procedure was identical to Experiment 2.

2.3.2 Results

This experiment was pre-registered at https://osf.io/89av6. A total of 456 participants completed the experiment. Eighty-two participants were excluded from the data analyses based on exclusion criteria (i.e., failed attention checks during the experiment, searched the answers to test questions, took notes during the lecture, did not watch the lectures fully). Seventeen participants were removed for completing the study twice. Sixty-two participants dropped out of the study. Fifty-two of the participants were in the freezing condition while ten of the

participants were in the no freezing condition. The difference in participant dropout by freezing condition was significant, X^2 (1, N = 360) = 23.17, p < .001. For the 52 participants in the freezing condition, 31 participants were in a front-loaded freezing condition, 10 participants were in a uniform freezing condition, and 11 participants were in a back-loaded freezing condition. The difference in participant dropout by freezing position condition was also significant, X^2 (2, N = 360) = 11.70, p = .003. Thus, there was clearly differential drop-out by condition. This needs to be considered when interpreting the reported results. Of the 62 participants that dropped out, 37 of the participants were in the slides condition while 25 of the participants were in the no slides condition. The difference in participant dropout by slides condition was not significant, X^2 (1, N = 360) = 1.61, p = .204. Excluded and dropout participants were replaced, thus target sample size was retained. The distribution of the 360 participants for each condition were as follows: 90 participants in the freezing condition and 90 in the no freezing condition. Of the 90 participants in the freezing condition, 31 were in the front-loaded freezing condition, 30 were in the uniform freezing condition, and 29 were in the back-loaded freezing condition. When asked about video issues, the majority of participants in the freezing condition (92%) responded that the video did contain freezing events. The majority of participants in the no freezing condition (97%) responded that the video did not contain freezing events.

The coding process for open response test questions was identical to Experiments 1 and 2. To examine the effect of freezing and slides on all variables, we conducted a 2 (freezing: freezing/no freezing) x 2 (slides: slides/no slides) between-subjects ANOVA. To examine the effect of freeze position on all variables, we conducted a 3 (freeze position: front-loaded/uniform/back-loaded) x 2 (slides: slides/no slides) between-subjects ANOVA restricted to

the freezing condition. Means and 95% confidence intervals for each dependent variable are shown in Table 8.

Table 8

Means and 95% CIs for dependent variables in Experiment 3.

	Freezing		No Freezing	
	Slides	No Slides	Slides	No Slides
Test Performance	52.22 [47.71–56.74]	41.31 [37.05–45.58]	54.19 [50.13–58.25]	46.21 [42.10–50.32]
Effort Invested	6.33 [6.00–6.67]	6.04 [5.70–6.39]	6.40 [6.10–6.70]	6.32 [6.00–6.65]
Effort Required	6.40 [6.05–6.75]	6.43 [6.07–6.80]	6.51 [6.17–6.85]	6.52 [6.11–6.93]
Overall MW	0.33 [0.27–0.40]	0.33 [0.26–0.39]	0.31 [0.24–0.37]	0.32 [0.26–0.38]
Intentional MW	0.05 [0.02–0.07]	0.07 [0.03–0.10]	0.03 [0.01–0.04]	0.04 [0.01–0.07]
Unintentional MW	0.29 [0.23–0.35]	0.26 [0.20–0.31]	0.28 [0.22–0.34]	0.28 [0.23–0.34]
JOL Before	57.87 [53.28–62.45]	53.61 [48.88–58.34]	58.00 [53.56–62.44]	56.41 [51.89–60.95]
JOL After	53.42 [47.91–58.94]	42.51 [37.23–47.79]	53.06 [47.70–58.41]	45.42 [39.62–51.23]
Liking (Experience)	3.69 [3.30–4.08]	2.92 [2.59–3.26]	3.86 [3.51–4.20]	3.59 [3.23–3.95]
Liking (Content)	4.37 [4.00–4.74]	3.87 [3.50–4.23]	4.10 [3.76–4.44]	4.03 [3.68–4.39]
Liking (Visuals)	3.33 [2.99–3.68]	2.06 [1.77–2.34]	3.01 [2.72–3.30]	2.43 [2.11–2.75]
Liking (Instructor)	4.72 [4.40–5.04]	4.32 [4.03–4.61]	4.80 [4.53–5.07]	4.60 [4.29–4.91]
Watch Again	3.52 [3.09–3.96]	2.32 [1.98–2.67]	3.41 [3.05–3.77]	2.58 [2.19–2.96]
Positive Affect	2.40 [1.94–2.31]	1.98 [1.82–2.13]	2.21 [2.05–2.36]	2.12 [1.94–2.31]
Negative Affect	1.22 [1.16–1.29]	1.25 [1.19–1.31]	1.14 [1.10–1.19]	1.21 [1.14–1.29]

Test Performance

There was no significant effect of freezing, $F(1, 356) = 2.59$, $p = .108$, $\eta^2_G = .007$, or interaction between freezing and slides, $F(1, 356) = 0.47$, $p = .493$, $\eta^2_G = .001$, on the proportion of correct answers to the test questions. There was a significant effect of slides, $F(1, 356) = 19.58$, $p < .001$, $\eta^2_G = .052$, such that participants in the slides condition performed better on the test compared to participants in the no slides condition. For the freeze position analyses, the pattern of results was similar (freeze position: $F(2, 174) = 0.64$, $p = .530$, $\eta^2_G = .007$; slides: $F(1, 174) = 12.05$, $p = .001$, $\eta^2_G = .065$; interaction: $F(2, 174) = 1.15$, $p = .318$, $\eta^2_G = .013$).

Effort

There was no significant effect of freezing, $F(1, 356) = 1.11$, $p = .293$, $\eta^2_G = .003$, slides, $F(1, 356) = 1.26$, $p = .263$, $\eta^2_G = .004$, or any interaction, $F(1, 356) = 0.42$, $p = .519$, $\eta^2_G = .001$, on invested effort. For the freeze position analyses, there was no significant effect of freeze position, $F(2, 174) = 0.01$, $p = .988$, $\eta^2_G < .001$, or slides, $F(1, 174) = 1.46$, $p = .229$, $\eta^2_G = .034$). There was a significant interaction between freeze position and slides, $F(2, 174) = 3.11$, $p = .047$, $\eta^2_G = .034$. This interaction appears to reflect a change in the ordering of conditions across the slides variable. In the slides condition, participants invested the most effort in the front-loaded freezing condition ($M = 6.73$) compared to the uniform freezing condition ($M = 6.13$) and back-loaded freezing condition ($M = 6.13$; $F(2, 87) = 1.42$, $p = .248$, $\eta^2_G = .032$), but in the no slides condition, participants invested the least effort in the front-loaded freezing condition ($M = 5.60$) compared to the uniform freezing condition ($M = 6.29$) and back-loaded freezing condition ($M = 6.24$; $F(2, 87) = 1.70$, $p = .189$, $\eta^2_G = .038$). There were no differences in invested effort across

the slides variable for participants in the uniform, $t(57) = 0.37$, $p = .715$, $d = 0.09$, and back-loaded conditions, $t(54) = 0.27$, $p = .789$, $d = 0.07$.

There was no significant effect of freezing, $F(1, 356) = 0.29$, $p = .590$, $\eta^2_G = .001$, slides, $F(1, 356) = 0.01$, $p = .905$, $\eta^2_G < .001$, or any interaction, $F(1, 356) = 0.004$, $p = .952$, $\eta^2_G < .001$, on required effort. For the freeze position analyses, there were no significant effects or interaction (freeze position: $F(2, 174) = 1.25$, $p = .288$, $\eta^2_G = .014$; slides: $F(1, 174) = 0.01$, $p = .907$, $\eta^2_G < .001$; interaction: $F(2, 174) = 1.25$, $p = .288$, $\eta^2_G = .014$).

Mind Wandering

There was no significant effect of freezing, $F(1, 356) = 0.26$, $p = .608$, $\eta^2_G = .001$, slides, $F(1, 356) = 0.01$, $p = .932$, $\eta^2_G < .001$, or any interaction, $F(1, 356) = 0.12$, $p = .732$, $\eta^2_G < .001$, on overall mind wandering. For the freeze position analyses, there were no significant effects or interaction (freeze position: $F(2, 174) = 0.17$, $p = .844$, $\eta^2_G = .002$; slides: $F(1, 174) = 0.03$, $p = .867$, $\eta^2_G < .001$; interaction: $F(2, 174) = 0.66$, $p = .520$, $\eta^2_G = .007$).

The overall pattern of results was also the same for intentional mind wandering (freezing: $F(1, 356) = 3.70$, $p = .055$, $\eta^2_G = .010$; slides: $F(1, 356) = 1.24$, $p = .267$, $\eta^2_G = .003$; interaction: $F(1, 356) = 0.09$, $p = .762$, $\eta^2_G < .001$; freeze position: $F(2, 174) = 0.17$, $p = .841$, $\eta^2_G = .002$; slides: $F(1, 174) = 0.77$, $p = .383$, $\eta^2_G = .004$; interaction: $F(2, 174) = 0.99$, $p = .373$, $\eta^2_G = .011$) and unintentional mind wandering (freezing: $F(1, 356) = 0.11$, $p = .737$, $\eta^2_G < .001$; slides: $F(1, 356) = 0.19$, $p = .666$, $\eta^2_G = .001$; interaction: $F(1, 356) = 0.28$, $p = .598$, $\eta^2_G = .001$; freeze position: $F(2, 174) = 0.21$, $p = .809$, $\eta^2_G = .002$; slides: $F(1, 174) = 0.42$, $p = .516$, $\eta^2_G = .002$; interaction: $F(2, 174) = 1.84$, $p = .163$, $\eta^2_G = .021$).

We also conducted a 2 (freeze density: dense freezing/no freezing) x 2 (slides: slides/no slides) mixed ANOVA on the proportion of reported mind wandering with freeze density as a within-subjects factor and slides as a between-subjects factor. Means and 95% CIs for the proportion of reported mind wandering are shown in Table 9. For overall mind wandering, there was no main effect of the slides condition, $F(1, 117) = 0.55$, $p = .461$, $\eta_G^2 = .003$, or interaction with freeze density, $F(1, 117) = 0.68$, $p = .411$, $\eta_G^2 = .002$. There was a significant main effect of freeze density, $F(1, 117) = 29.83$, $p < .001$, $\eta_G^2 = .094$, such that participants reported more overall mind wandering in the dense freezing condition ($M = 0.51$) compared to the no freezing condition ($M = 0.22$). This pattern of results was the same for intentional mind wandering, (slides: $F(1, 117) = 0.03$, $p = .860$, $\eta_G^2 < .001$; freeze density: $F(1, 117) = 14.67$, $p < .001$, $\eta_G^2 = .049$; slides x freeze density: $F(1, 117) = 0.08$, $p = .774$, $\eta_G^2 < .001$), and unintentional mind wandering, (slides: $F(1, 117) = 0.77$, $p = .382$, $\eta_G^2 = .004$; freeze density: $F(1, 117) = 10.81$, $p = .001$, $\eta_G^2 = .036$; slides x freeze density: $F(1, 117) = 0.47$, $p = .495$, $\eta_G^2 = .002$).

Table 9

Means and 95% CIs for the proportion of reported mind wandering during the video lecture across freeze density (dense freezing, no freezing) and the slides condition (slides, no slides) in Experiment 3.

	Slides		No Slides	
	Dense Freezing	No Freezing	Dense Freezing	No Freezing
Overall MW	0.47 [0.34–0.60]	0.22 [0.11–0.32]	0.56 [0.43–0.69]	0.22 [0.11–0.33]
Intentional MW	0.15 [0.06–0.24]	0.03 [-0.01–0.08]	0.15 [0.06–0.25]	0.02 [-0.02–0.05]
Unintentional MW	0.31 [0.20–0.44]	0.18 [0.08–0.28]	0.41 [0.28–0.54]	0.20 [0.10–0.31]

Metacognition

There was no significant effect of freezing, $F(1, 356) = 0.41$, $p = .523$, $\eta^2_G = .001$, slides, $F(1, 356) = 1.61$, $p = .206$, $\eta^2_G = .004$, or any interaction, $F(1, 356) = 0.34$, $p = .561$, $\eta^2_G = .001$, on predicted scores before test. For the freeze position analyses, there were no significant effects or interaction (freeze position: $F(2, 174) = 1.86$, $p = .159$, $\eta^2_G = .021$; slides: $F(1, 174) = 1.68$, $p = .197$, $\eta^2_G = .010$; interaction: $F(2, 174) = 0.45$, $p = .638$, $\eta^2_G = .005$).

There was no significant effect of freezing, $F(1, 356) = 0.21$, $p = .646$, $\eta^2_G = .001$, or interaction with slides, $F(1, 356) = 0.35$, $p = .554$, $\eta^2_G = .001$, on predicted scores after test. There was a significant effect of slides, $F(1, 356) = 11.26$, $p = .001$, $\eta^2_G = .031$, such that participants in the slides condition reported a higher predicted score after test compared to participants in the no slides condition. For the freeze position analyses, the pattern of results was similar (freeze position: $F(2, 174) = 0.64$, $p = .530$, $\eta^2_G = .007$; slides: $F(1, 174) = 12.05$, $p = .001$, $\eta^2_G = .065$; interaction: $F(2, 174) = 1.15$, $p = .318$, $\eta^2_G = .013$).

For predicted scores before test, there was a positive correlation between prospective predicted scores and actual test performance in the freezing, $r(178) = 0.41$, $p < .001$, and no freezing condition, $r(178) = 0.39$, $p < .001$.

For predicted scores after test, there was a positive correlation between retrospective predicted scores and actual test performance in the freezing, $r(178) = 0.62$, $p < .001$, and no freezing condition, $r(178) = 0.65$, $p < .001$.

Affect

Liking. There was a significant effect of freezing on liking the general experience of the lecture, $F(1, 356) = 5.34$, $p = .021$, $\eta^2_G = .015$, such that participants in the freezing condition

reported lower liking for the general experience of the lecture compared to participants in the no freezing condition. There was also a significant effect of slides, $F(1, 356) = 8.20$, $p = .004$, $\eta^2_G = .023$, such that participants in the slides condition reported higher liking for the general experience of the lecture compared to participants in the no slides condition. There was no significant interaction between freezing or slides, $F(1, 356) = 1.92$, $p = .167$, $\eta^2_G = .005$. In the freeze position analyses, there was no significant effect of freeze position on liking the general experience of the lecture, $F(2, 174) = 0.20$, $p = .821$, $\eta^2_G = .002$, or any interaction with slides, $F(2, 174) = 2.64$, $p = .074$, $\eta^2_G = .029$.

There was no significant effect of freezing, $F(1, 356) = 0.08$, $p = .781$, $\eta^2_G < .001$, slides, $F(1, 356) = 2.48$, $p = .116$, $\eta^2_G = .007$, or any interaction, $F(1, 356) = 1.45$, $p = .230$, $\eta^2_G = .004$, on liking the content of the lecture. For the freeze position analyses, there were no significant effects or interaction (freeze position: $F(2, 174) = 0.59$, $p = .555$, $\eta^2_G = .007$; slides: $F(1, 174) = 3.75$, $p = .054$, $\eta^2_G = .021$; interaction: $F(2, 174) = 1.72$, $p = .182$, $\eta^2_G = .019$).

There was no significant effect of freezing, $F(1, 356) = 0.03$, $p = .859$, $\eta^2_G < .001$, on liking the visuals of the lecture. There was a significant effect of slides, $F(1, 356) = 35.19$, $p < .001$, $\eta^2_G = .090$, such that participants in the slides condition reported higher liking for the visuals of the lecture compared to participants in the no slides condition. There was also a significant interaction, $F(1, 356) = 5.01$, $p = .026$, $\eta^2_G = .014$. This interaction reflects a reversal in the direction of the freezing effect across the slides condition. With slides, participants in the freezing condition preferred the visuals slightly more ($M = 3.33$) compared to participants in the no freezing condition ($M = 3.01$), $t(173) = 1.42$, $p = .157$, $d = 0.21$. With no slides, participants in the no freezing condition preferred the visuals slightly more ($M = 2.43$) compared to participants in the freezing condition ($M = 2.06$), $t(176) = 1.75$, $p = .082$, $d = 0.26$.

In the freeze position analyses, the effect of slides was the only significant effect (freeze position: $F(2, 174) = 0.99$, $p = .374$, $\eta^2_G = .011$; slides: $F(1, 174) = 32.82$, $p < .001$, $\eta^2_G = .159$; interaction: $F(2, 174) = 2.53$, $p = .083$, $\eta^2_G = .028$).

There was no significant effect of freezing, $F(1, 356) = 1.40$, $p = .238$, $\eta^2_G = .004$, or any interaction with slides, $F(1, 356) = 0.44$, $p = .507$, $\eta^2_G = .001$, on liking the instructor of the lecture. There was a significant effect of slides, $F(1, 356) = 3.97$, $p = .047$, $\eta^2_G = .011$, such that participants in the slides condition reported higher liking for the instructor of the lecture compared to participants in the no slides condition. In the freeze position analyses, there was no significant effect of freeze position, $F(2, 174) = 1.49$, $p = .229$, $\eta^2_G = .017$, or slides, $F(1, 174) = 3.55$, $p = .061$, $\eta^2_G = .020$. There was a significant interaction, $F(2, 174) = 4.10$, $p = .018$, $\eta^2_G = .045$). Follow-up analyses revealed that in the slides condition, there were significant differences between freeze position conditions, $F(2, 87) = 3.89$, $p = .024$, $\eta^2_G = .082$. Post-hoc comparisons revealed no differences between the front-loaded freezing and the back-loaded freezing conditions, $t(57) = 1.48$, $p = .145$, $d = 0.38$, and between the front-loaded freezing and the uniform freezing conditions, $t(57) = 1.31$, $p = .196$, $d = 0.33$. There was a significant difference between the uniform freezing and the back-loaded freezing conditions, $t(57) = 2.80$, $p = .007$, $d = 0.72$, such that participants in the back-loaded freezing condition liked the instructor more ($M = 3.83$) than participants in uniform freezing condition ($M = 2.83$). In the no slides condition, there were no significant differences between the freeze position conditions, $F(2, 87) = 1.53$, $p = .221$, $\eta^2_G = .034$.

Likelihood of watching the lecture in the same format. There was no significant effect of freezing, $F(1, 356) = 0.14$, $p = .708$, $\eta^2_G < .001$, or any interaction with slides, $F(1, 356) = 0.90$, $p = .342$, $\eta^2_G = .003$, on the likelihood of watching the lecture in the same format. There was a

significant effect of slides, $F(1, 356) = 27.79$, $p < .001$, $\eta^2_G = .072$, such that participants in the slides condition were more likely to watch the lecture in the same format compared to participants in the no slides condition. This pattern of results was the same for the freeze position analyses (freeze position: $F(2, 174) = 2.06$, $p = .130$, $\eta^2_G = .023$; slides: $F(1, 174) = 18.99$, $p < .001$, $\eta^2_G = .098$; interaction: $F(2, 174) = 2.42$, $p = .092$, $\eta^2_G = .027$).

PANAS. There was no significant effect of freezing, $F(1, 356) = 0.08$, $p = .784$, $\eta^2_G < .001$, on positive affect. There was a significant effect of slides, $F(1, 356) = 8.53$, $p = .004$, $\eta^2_G = .023$, such that participants in the slides condition reported higher positive affect compared to participants in the no slides condition. There was also a significant interaction, $F(1, 356) = 3.95$, $p = .048$, $\eta^2_G = .011$. This interaction reflects a reversal in the direction of the freezing effect across the slides condition. With slides, participants in the freezing condition reported slightly higher positive affect ($M = 2.40$) compared to participants in the no freezing condition ($M = 2.21$), $t(169) = 1.57$, $p = .118$, $d = 0.23$. With no slides, participants in the no freezing condition reported slightly higher positive affect ($M = 2.12$) compared to participants in the freezing condition ($M = 1.98$), $t(173) = 1.23$, $p = .219$, $d = 0.18$. In the freeze position analyses, the effect of slides was the only significant effect (freeze position: $F(2, 174) = 1.22$, $p = .299$, $\eta^2_G = .014$; slides: $F(1, 174) = 11.75$, $p = .001$, $\eta^2_G = .063$; interaction: $F(2, 174) = 1.62$, $p = .201$, $\eta^2_G = .018$). There was no significant effect of freezing, $F(1, 356) = 3.48$, $p = .063$, $\eta^2_G = .010$, slides, $F(1, 356) = 2.73$, $p = .100$, $\eta^2_G = .008$, or any interaction, $F(1, 356) = 0.48$, $p = .489$, $\eta^2_G = .001$, on negative affect. This pattern of results was the same for the freeze position analyses (freeze position: $F(2, 174) = 0.49$, $p = .613$, $\eta^2_G = .006$; slides: $F(1, 174) = 0.49$, $p = .483$, $\eta^2_G = .003$; interaction: $F(2, 174) = 1.71$, $p = .183$, $\eta^2_G = .019$).

Bivariate Correlations

We examined the bivariate relations between the various dependent variables, which are displayed in Table 10. Again, significance values should be interpreted cautiously given the number of correlations. As with the previous experiments, the correlations were generally consistent (e.g., test performance was positively correlated with invested effort and with liking and affect, and negatively related to required effort, negative affect and mind wandering).

Table 10

Bivariate correlation table for Experiment 3

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Test Performance														
2. Effort Invested	.17**													
3. Effort Required	-.18**	.34**												
4. Overall MW	-.25**	-.55**	.01											
5. Intentional MW	-.16**	-.29**	-.06	.45**										
6. Unintentional MW	-.21**	-.48**	.04	.91**	.03									
7. JOL Before	.40**	.30**	-.23**	-.38**	-.18**	-.34**								
8. JOL After	.63**	.29**	-.19**	-.35**	-.21**	-.29**	.72**							
9. Liking (Experience)	.21**	.37**	-.01	-.37**	-.13*	-.35**	.41**	.33**						
10. Liking (Content)	.33**	.41**	-.11*	-.43**	-.16**	-.41**	.52**	.45**	.72**					
11. Liking (Visuals)	.10*	.31**	.02	-.25**	-.12*	-.22**	.18**	.12*	.53**	.40**				
12. Liking (Instructor)	.05	.31**	-.11*	-.24**	-.07	-.24**	.23**	.17**	.52**	.44**	.48**			
13. Watch Again	.21	.38**	-.03	-.38**	-.16**	-.35**	.32**	.26**	.69**	.62**	.69**	.52**		
14. Positive Affect	.22	.63**	.16	-.51**	-.18**	-.48**	.33	.30**	.61**	.60**	.52**	.48**	.66**	
15. Negative Affect	-.10*	-.06	.08	.22**	.12*	.19**	-.25**	-.21**	-.29**	-.23**	-.11*	-.20**	-.20**	-.14*

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

2.3.3 Discussion

As in Experiments 1 and 2, there were no significant effects of freezing or the freeze position conditions on test performance, positive and negative affect, and effort. There was no overall effect of freezing on overall, intentional, and unintentional mind wandering. However, with respect to the effect of freeze density, there was a significant main effect of freeze density. Participants reported more overall mind wandering in the dense freezing condition compared to the no freezing condition. These results suggest that dense freezing periods might cause an increase in the likelihood of disengagement. There were no significant effects of freezing on predicted scores before and after test. There were no effects of freezing on liking the content, visuals, and instructor. There was a significant effect of freezing on liking the general experience of the lecture however, unlike Experiment 2, participants in the freezing condition reported lower ratings for liking the general experience of the lecture compared to participants in the no freezing condition. There was no effect of freezing on the likelihood of watching the lecture in the same format. While the presence of slides did not appear to moderate the effect of freezing significantly, there were significant effects of slides on test performance, predicted scores after test, and overall affect, such that participants who had slides did better on the test, had higher predicted scores after test, reported higher liking for the general experience, visuals, and instructor of the lecture as well as reported higher positive affect.

Freezing also impacted the likelihood that participants would drop out. Of the 62 participants who dropped out of the study, 52 participants were in a freezing condition. In addition, 31 participants were in the front-loaded freezing condition. As a result, our data only included individuals who were willing to push through the freezing, creating a potentially critical difference between that group and the no freezing group. This needs to be considered when

interpreting the results of Experiment 3. The imbalance present wherein participant drop-out was particularly bad in the front-loaded freezing condition might be consistent with the mind-wandering results wherein those that remained spent a disproportionate amount of time mind wandering during the dense freezing periods (i.e., some participants appeared to disengage mentally and others disengage completely).

2.4 Experiment 4

The longer duration freezing events in Experiment 3 did appear to lead to negative impacts (e.g., on mind wandering, liking). However, in Experiment 3, freezing led to differential dropout making interpreting differences across conditions difficult. As such, we sought to replicate our results thus far with an in-person sample where drop-out might be less likely. We used the same stimuli from both Experiments 2 and 3 to compare across three conditions: no freezing events, five-second freezing events, and fifteen-second freezing events. In the freezing conditions, the freezing events were distributed either in the first three minutes of the video (front-loaded), throughout the video (uniform), or the last three minutes of the video (back-loaded). We also added four questions to assess participants' ratings of the video quality of the lecture as well as how the video quality affected their ability to understand, their ability to pay attention to, and their enjoyment of the lecture.

2.4.1 Methods

Participants

An a-priori sample size was determined to detect a medium effect of freezing on all the dependent variables (Cohen's d = .50) with a power of .95 (alpha = .05, two-tailed) in a 3 x 3 between-subjects ANOVA. Participants (N = 273; reported gender: 65% female, 0.73%

unreported; M_{age} = 22.13 years) were recruited from the University of Waterloo (n = 209) and from the online platform Prolific (n = 64). To conduct this experiment with an in-person sample, participants from the university completed the experiment while a researcher was present in the room. To emulate this environment, online Prolific participants were informed that they would be participating in a synchronous study in which they would be completing the experiment online while a researcher was present with them on a video call (i.e., Zoom). Participants from the university received course credit and Prolific participants received GBP £4.00 as remuneration for the completion of the study. All participants reported their highest level of education (0.37% no formal education, 76% high school diploma, 4% college diploma, 11% bachelor's degree, 7% master's degree, 1% doctorate degree), the number of online courses they have taken (M = 8.93), and their English proficiency in reading (89% high proficiency, 10% moderate proficiency) and listening (91% high proficiency, 9% moderate proficiency).

Stimuli

The video lectures were similar to those used in previous experiments except there were three video conditions: video lecture with no freezing events, video lecture with five-second freezing events, video lecture with fifteen-second freezing events. All video lectures contained visual content (i.e., lecture slides with information).

Measures

Participants completed the same demographic questionnaire, mind wandering probes, content test questions and rating measures (i.e., liking, likelihood of watching the lecture in the same format, affect, and effort) that were used in Experiment 3. Participants were also asked four questions on video quality. On a scale from 1–5 (1 = Very bad, 5 = Very good), participants rated the overall video quality of the lecture. On a scale from 1–7 (1 = Very negative effect, 7 =

Very positive effect), participants rated how much the video quality of the lecture affected their ability to understand the material, their ability to pay attention to the material, and their overall enjoyment of the material.

Procedure

The procedure was identical to Experiment 3.

2.4.2 Results

This experiment was pre-registered at https://osf.io/egqd5. A total of 312 participants completed the experiment. Thirty-eight participants were excluded from the data analyses based on exclusion criteria (i.e., failed attention checks during the experiment, searched the answers to test questions, took notes during the lecture, did not watch the lectures fully). Excluded participants were replaced, thus target sample size was retained. Participants who did not complete the metacognitive ratings were removed from the metacognitive rating analyses (1 participant). No participants dropped out of the study. The distribution of the 273 participants for each condition were as follows: 92 participants saw a no freezing condition, 91 participants saw a five-second freezing condition, and 90 participants saw a fifteen-second freezing condition. Of the 90 participants in the freezing condition, 31 saw a front-loaded freezing condition, 30 saw a uniform freezing condition, and 29 saw a back-loaded freezing condition. When asked about video issues, the majority of participants in the five-second freezing condition (78%) responded that the video did contain freezing events. The majority of participants in the fifteen-second freezing condition (84%) responded that the video did contain freezing events. The majority of participants in the no freezing condition (78%) responded that the video did not contain freezing events.

The coding process for open response test questions was identical to Experiments 1, 2, and 3. To examine the effect of freezing on all variables, we conducted a 3 (freezing: no freezing/five/fifteen) x 3 (freeze position: front-loaded/uniform/back-loaded) between-subjects ANOVA, with freeze position restricted to the two freezing conditions. Means and 95% confidence intervals for each dependent variable are shown in Table 11.

Table 11

Means and 95% CIs for dependent variables in Experiment 4.

	No Freezing			Five-Second			Fifteen-Second		
	Front-loaded	Uniform	Back-loaded	Front-loaded	Uniform	Back-loaded	Front-loaded	Uniform	Back-loaded
Test Performance	59.42 [50.34–68.49]	56.60 [48.48–64.62]	60.61 [52.28–68.93]	56.25 [47.70–64.80]	52.35 [42.54–62.16]	55.76 [45.92–65.59]	62.38 [53.53–71.23]	60.61 [53.40–67.81]	59.09 [50.44–67.74]
Effort Invested	6.00 [5.47–6.53]	5.58 [5.04–6.12]	5.76 [5.14–6.37]	5.63 [5.01–6.24]	5.52 [4.81–6.23]	5.63 [4.98–6.29]	5.90 [5.38–6.41]	5.40 [4.73–6.07]	5.81 [5.34–6.28]
Effort Required	6.21 [5.60–6.82]	5.81 [5.17–6.44]	5.79 [5.11–6.47]	5.59 [4.84–6.35]	5.48 [4.68–6.29]	5.73 [5.09–6.38]	5.66 [4.86–6.45]	5.33 [4.53–6.14]	5.59 [4.93–6.25]
Overall MW	0.50 [0.38–0.62]	0.48 [0.36–0.61]	0.47 [0.38–0.56]	0.52 [0.42–0.62]	0.60 [0.51–0.70]	0.49 [0.38–0.61]	0.45 [0.34–0.55]	0.62 [0.50–0.73]	0.53 [0.42–0.64]
Intentional MW	0.10 [0.03–0.17]	0.13 [0.06–0.20]	0.11 [0.05–0.16]	0.20 [0.11–0.29]	0.15 [0.09–0.22]	0.08 [0.03–0.13]	0.10 [0.04–0.16]	0.17 [0.09–0.24]	0.21 [0.14–0.28]
Unintentional MW	0.40 [0.30–0.51]	0.35 [0.25–0.46]	0.36 [0.28–0.44]	0.31 [0.22–0.40]	0.45 [0.34–0.55]	0.41 [0.31–0.51]	0.34 [0.25–0.44]	0.45 [0.34–0.56]	0.32 [0.23–0.41]
JOL Before	63.96 [57.46–70.47]	62.10 [56.24–67.95]	58.94 [51.91–65.97]	62.91 [56.55–69.26]	57.07 [48.17–65.97]	61.93 [53.34–70.53]	60.90 [53.61–68.19]	62.50 [54.66–70.34]	63.13 [56.91–69.35]
JOL After	59.29 [49.16–69.41]	60.32 [52.73–67.91]	57.58 [48.16–66.99]	57.89 [48.75–67.03]	53.55 [42.37–64.73]	62.03 [52.41–71.66]	64.83 [55.83–73.83]	63.43 [52.99–73.88]	56.77 [47.29–66.26]
Liking (Experience)	4.25 [3.86–4.64]	4.52 [3.99–5.04]	3.76 [3.20–4.31]	3.72 [3.18–4.26]	3.41 [2.86–3.97]	3.50 [2.98–4.02]	3.62 [3.08–4.16]	2.73 [2.18–3.29]	3.38 [2.80–3.95]
Liking (Content)	4.86 [4.34–5.37]	4.74 [4.21–5.28]	3.97 [3.30–4.64]	4.59 [4.04–5.15]	4.21 [3.51–4.91]	4.80 [4.21–5.39]	4.31 [3.76–4.86]	4.57 [3.89–5.24]	4.65 [4.18–5.14]
Liking (Visuals)	3.14 [2.55–3.74]	3.16 [2.49–3.83]	2.91 [2.31–3.51]	3.06 [2.47–3.66]	2.90 [2.20–3.59]	3.00 [2.46–3.54]	2.83 [2.29–3.37]	3.23 [2.68–3.79]	2.58 [2.19–3.00]
Liking (Instructor)	4.43 [3.96–4.89]	4.45 [3.93–4.97]	4.24 [3.73–4.76]	4.63 [4.03–5.22]	4.59 [4.17–5.00]	4.33 [3.84–4.83]	4.59 [4.15–5.02]	4.27 [3.76–4.78]	4.19 [3.69–4.75]

	No Freezing			Five-Second			Fifteen-Second		
	Front-loaded	Uniform	Back-loaded	Front-loaded	Uniform	Back-loaded	Front-loaded	Uniform	Back-loaded
Watch Again	3.14 [2.42–3.86]	3.42 [2.69–4.15]	2.64 [2.03–3.24]	3.00 [2.37–3.63]	2.62 [1.95–3.29]	3.13 [2.48–3.79]	2.45 [1.94–2.95]	2.57 [1.88–3.25]	2.77 [2.24–3.45]
Positive Affect	1.99 [1.75–2.22]	2.15 [1.90–2.40]	2.04 [1.72–2.36]	2.12 [1.86–2.39]	2.09 [1.78–2.40]	2.12 [1.87–2.37]	2.00 [1.76–2.24]	1.91 [1.67–2.15]	1.94 [1.69–2.19]
Negative Affect	1.16 [1.09–1.23]	1.24 [1.14–1.34]	1.26 [1.09–1.44]	1.42 [1.23–1.61]	1.32 [1.15–1.49]	1.34 [1.19–1.50]	1.29 [1.17–1.40]	1.31 [1.19–1.44]	1.48 [1.34–1.63]
Video Quality (VQ)	3.71 [3.32–4.11]	3.61 [3.23–4.00]	3.48 [3.15–3.82]	2.19 [1.89–2.48]	2.14 [1.79–2.49]	2.47 [2.18–2.76]	1.93 [1.69–2.18]	1.97 [1.68–2.25]	2.06 [1.82–2.30]
VQ Understanding	4.25 [3.68–4.81]	4.19 [3.65–4.74]	3.88 [3.37–4.39]	2.72 [2.30–3.14]	3.14 [2.71–3.57]	3.07 [2.64–3.49]	2.62 [2.14–3.10]	3.20 [2.63–3.77]	2.72 [2.18–3.26]
VQ Attention	3.96 [3.36–4.57]	3.77 [3.24–4.31]	3.64 [3.11–4.17]	2.38 [1.84–2.91]	2.52 [2.09–2.94]	2.67 [2.27–3.06]	2.07 [1.64–2.50]	2.27 [2.70–2.84]	2.28 [1.75–2.82]
VQ Enjoyment	4.04 [3.40–4.67]	3.74 [3.17–4.32]	3.48 [2.97–4.00]	2.19 [1.69–2.68]	2.38 [1.92–2.84]	2.17 [1.79–2.55]	1.93 [1.59–2.27]	2.17 [1.66–2.68]	2.06 [1.69–2.44]

Test Performance

There was no significant effect of freezing, $F(2, 265) = 1.50$, $p = .225$, $\eta^2_G = .011$, freeze position, $F(2, 265) = 0.34$, $p = .712$, $\eta^2_G = .003$, or interaction, $F(4, 265) = 0.15$, $p = .965$, $\eta^2_G = .002$, on the proportion of correct answers to the content test questions.

Effort

There was no significant effect of freezing, $F(2, 265) = 0.31$, $p = .731$, $\eta^2_G = .002$, freeze position, $F(2, 265) = 1.06$, $p = .349$, $\eta^2_G = .008$, or interaction, $F(4, 265) = 0.16$, $p = .957$, $\eta^2_G = .002$, on invested effort. This pattern of results was the same for required effort (freezing: $F(2, 265) = 1.16$, $p = .314$, $\eta^2_G = .009$; freeze position: $F(2, 265) = 0.48$, $p = .621$, $\eta^2_G = .004$; interaction: $F(4, 265) = 0.18$, $p = .948$, $\eta^2_G = .003$).

Mind Wandering

There was no significant effect of freezing, $F(2, 265) = 0.91$, $p = .404$, $\eta^2_G = .007$, freeze position, $F(2, 265) = 2.05$, $p = .131$, $\eta^2_G = .015$, or interaction, $F(4, 265) = 0.89$, $p = .471$, $\eta^2_G = .013$, on overall mind wandering. This pattern of results was the same for unintentional mind wandering (freezing: $F(2, 265) = 0.13$, $p = .880$, $\eta^2_G = .001$; freeze position: $F(2, 265) = 1.54$, $p = .216$, $\eta^2_G = .012$; interaction: $F(4, 265) = 1.45$, $p = .219$, $\eta^2_G = .021$).

There was no significant effect of freezing, $F(2, 265) = 1.69$, $p = .186$, $\eta^2_G = .013$, or freeze position, $F(2, 265) = 0.23$, $p = .799$, $\eta^2_G = .002$, on intentional mind wandering. There was a significant interaction, $F(4, 265) = 2.87$, $p = .023$, $\eta^2_G = .042$. Follow-up analyses revealed that in the back-loaded condition, there were significant differences between the duration conditions, $F(2, 92) = 5.25$, $p = .007$, $\eta^2_G = .102$. Participants in the fifteen-second freezing condition ($M = 0.21$) reported more intentional mind wandering compared to participants in the no freezing

condition (M = 0.11), $t(61) = 2.36$, $p = .022$, $d = 0.59$, and participants in the five-second freezing condition (M = 0.08), $t(56) = 3.03$, $p = .004$, $d = 0.76$. There were no significant differences between the no freezing and five-second conditions, $t(60) = 0.60$, $p = .554$, $d = 0.15$. There were no significant differences between the duration conditions in the front-loaded, $F(2, 86) = 2.61$, $p = .079$, $\eta_G^2 = .057$, and uniform conditions, $F(2, 87) = 0.31$, $p = .735$, $\eta_G^2 = .007$.

Similar to previous experiments, we investigated the effect of freeze density. Means and 95% CIs for the proportion of reported mind wandering are shown in Table 12. For overall mind wandering, there was no main effect of freeze duration, $F(2, 181) = 0.42$, $p = .657$, $\eta_G^2 = .003$. There was a significant main effect of freeze density, $F(1, 181) = 11.37$, $p < .001$, $\eta_G^2 = .025$, such that participants reported more overall mind wandering in the dense freezing condition (M = 0.65) compared to the no freezing condition (M = 0.50). There was also a significant interaction between freeze duration and freeze density, $F(2, 181) = 10.97$, $p < .001$, $\eta_G^2 = .047$. This interaction reflects a change in the effect of freeze density across the freeze duration conditions.

There was a significant difference in the effect of freeze density between the no freezing and five-second freezing conditions, $t(121) = 3.03$, $p = .003$, $d = 0.55$. In the five-second freezing condition, participants reported more overall mind wandering in the dense freezing condition (M = 0.69) compared to the no freezing condition (M = 0.48). In the no freezing condition, there was no difference in the proportion of mind wandering between the dense freezing (M = 0.48) and no freezing conditions (M = 0.61). There was also a significant difference in the effect of freeze density between the no freezing and fifteen-second freezing conditions, $t(120) = 4.67$, $p < .001$, $d = 0.85$. In the fifteen-second freezing condition, participants reported more overall mind wandering in the dense freezing condition (M = 0.78) compared to the no freezing condition (M

= 0.40). There was no significant difference in the effect of freeze density between the five-second and fifteen-second freezing conditions, $t(121) = 1.53$, $p = .129$, $d = 0.28$.

For intentional mind wandering, there was a main effect of freeze duration, $F(2, 181) = 3.76$, $p = .025$, $\eta_G^2 = .022$. Participants in the fifteen-second freezing condition ($M = 0.24$) reported more intentional mind wandering compared to participants in the no freezing condition ($M = 0.11$), $t(220) = 2.74$, $p = .007$, $d = 0.35$. There were no significant differences in intentional mind wandering between participants in the fifteen-second freezing ($M = 0.24$) and five-second freezing conditions ($M = 0.16$), $t(238) = 1.50$, $p = .135$, $d = 0.19$, as well as between participants in the no freezing and five-second freezing conditions, $t(238) = 1.26$, $p = .209$, $d = 0.16$. There was a main effect of freeze density, $F(1, 181) = 16.00$, $p < .001$, $\eta_G^2 = .038$, such that participants reported more intentional mind wandering in the dense freezing condition ($M = 0.24$) compared to the no freezing condition ($M = 0.10$). There was also a significant interaction, $F(2, 181) = 10.30$, $p < .001$, $\eta_G^2 = .049$. This interaction reflects a change in the effect of freeze density across the freeze duration conditions. There was a significant difference in the effect of freeze density between the no freezing and five-second freezing conditions, $t(121) = 2.23$, $p = .028$, $d = 0.40$. In the five-second freezing condition, participants reported more intentional mind wandering in the dense freezing condition ($M = 0.23$) compared to the no freezing condition ($M = 0.10$). In the no freezing condition, there was no difference in the proportion of intentional mind wandering between the dense freezing ($M = 0.08$) and no freezing conditions ($M = 0.13$). There was also a significant difference in the effect of freeze density between the no freezing and fifteen-second freezing conditions, $t(113) = 4.45$, $p < .001$, $d = 0.81$. In the fifteen-second freezing condition, participants reported more intentional mind wandering in the dense freezing condition ($M = 0.41$) compared to the no freezing condition ($M = 0.07$). There was also a significant difference in the

effect of freeze density between the five-second and fifteen-second freezing conditions, $t(117) = 2.36$, $p = .020$, $d = 0.43$. For unintentional mind wandering, there was no significant effect of freeze duration, $F(2, 181) = 0.77$, $p = .463$, $\eta_G^2 = .005$, freeze density, $F(1, 181) = 0.45$, $p = .832$, $\eta_G^2 < .001$, or interaction, $F(2, 181) = 0.96$, $p = .386$, $\eta_G^2 = .005$.

Metacognition

There was no significant effect of freezing, $F(2, 264) = 0.15$, $p = .863$, $\eta_G^2 = .001$, freeze position, $F(2, 264) = 0.25$, $p = .780$, $\eta_G^2 = .002$, or interaction, $F(4, 264) = 0.58$, $p = .680$, $\eta_G^2 = .009$, on predicted scores before test. This pattern of results was the same for predicted scores after test (freezing: $F(2, 264) = 0.53$, $p = .592$, $\eta_G^2 = .004$; freeze position: $F(2, 264) = 0.14$, $p = .872$, $\eta_G^2 = .001$; interaction: $F(4, 264) = 0.80$, $p = .527$, $\eta_G^2 = .012$).

For predicted scores before test, there was a positive correlation between prospective predicted scores and actual test performance in the no freezing, $r(90) = 0.43$, $p < .001$, five-second freezing, $r(90) = 0.64$, $p < .001$, and fifteen-second freezing condition, $r(90) = 0.53$, $p < .001$.

For predicted scores after test, there was a positive correlation between retrospective predicted scores and actual test performance in the no freezing, $r(90) = 0.63$, $p < .001$, five-second freezing, $r(90) = 0.74$, $p < .001$, and fifteen-second freezing condition, $r(90) = 0.68$, $p < .001$.

Table 12

Means and 95% CIs for the proportion of reported mind wandering during the video lecture across freeze density (dense freezing, no freezing) and the freeze duration condition (no freezing, five seconds, fifteen seconds) in Experiment 4.

	No Freezing		Five-Second		Fifteen-Second	
	Dense Freezing	No Freezing	Dense Freezing	No Freezing	Dense Freezing	No Freezing
Overall MW	0.48 [0.35–0.60]	0.61 [0.48–0.73]	0.69 [0.58–0.81]	0.48 [0.136–0.61]	0.79 [0.68–0.89]	0.41 [0.28–0.54]
Intentional MW	0.08 [0.01–0.15]	0.13 [0.04–0.22]	0.23 [0.12–0.33]	0.10 [0.02–0.17]	0.41 [0.28–0.54]	0.07 [0.002–0.13]
Unintentional MW	0.39 [0.27–0.52]	0.48 [0.35–0.60]	0.47 [0.34–0.60]	0.39 [0.26–0.51]	0.38 [0.25–0.50]	0.34 [0.22–0.47]

Affect

Liking. There was a significant effect of freezing on liking the general experience of the lecture, $F(2, 265) = 9.90$, $p < .001$, $\eta^2_G = .070$, such that participants in the five-second freezing condition, $t(181) = 2.93$, $p = .004$, $d = 0.43$, and the fifteen-second freezing condition, $t(179) = 4.25$, $p < .001$, $d = 0.63$, reported lower liking for the general experience of the lecture compared to participants in the no freezing condition. There was no difference between participants in the five-second and fifteen-second freezing conditions, $t(179) = 1.40$, $p = .165$, $d = 0.21$. There was no significant effect of freeze position, $F(2, 265) = 1.42$, $p = .244$, $\eta^2_G = .011$, or interaction with freezing, $F(4, 265) = 2.11$, $p = .080$, $\eta^2_G = .031$.

There was no significant effect of freezing, $F(2, 265) = 0.01$, $p = .996$, $\eta^2_G < .001$, freeze position, $F(2, 265) = 0.12$, $p = .887$, $\eta^2_G = .001$, or interaction, $F(4, 265) = 2.09$, $p = .082$, $\eta^2_G = .031$, on liking the content of the lecture. This pattern of results was the same for liking the visuals of the lecture (freezing: $F(2, 265) = 0.32$, $p = .725$, $\eta^2_G = .002$; freeze position: $F(2, 265) = 0.68$, $p = .509$, $\eta^2_G = .005$; interaction: $F(4, 265) = 0.47$, $p = .756$, $\eta^2_G = .007$) and liking the instructor of the lecture (freezing: $F(2, 265) = 0.37$, $p = .693$, $\eta^2_G = .003$; freeze position: $F(2, 265) = 1.00$, $p = .370$, $\eta^2_G = .007$; interaction: $F(4, 265) = 0.14$, $p = .967$, $\eta^2_G = .002$).

Likelihood of watching the lecture in the same format. There was no significant effect of freezing, $F(2, 265) = 1.54$, $p = .217$, $\eta^2_G = .011$, freeze position, $F(2, 265) < 0.01$, $p = 1$, $\eta^2_G < .001$, or interaction, $F(4, 265) = 1.37$, $p = .245$, $\eta^2_G = .020$, on the likelihood of watching the lecture in the same format.

PANAS. There was no significant effect of freezing, $F(2, 265) = 1.17$, $p = .311$, $\eta^2_G = .009$, freeze position, $F(2, 265) = 0.02$, $p = .984$, $\eta^2_G < .001$, or interaction, $F(4, 265) = 0.26$, $p = .903$,

η^2_G = .004, on positive affect. There was a significant effect of freezing on negative affect, F(2, 265) = 3.82, p = .023, η^2_G = .028, such that participants in the five-second freezing condition, t(168) = 2.30, p = .022, d = 0.34, and the fifteen-second freezing condition, t(181) = 2.64, p = .009, d = 0.39, reported higher negative affect compared to participants in the no freezing condition. There was no difference between participants in the five-second and fifteen-second freezing conditions, t(171) = 0, p = 1, d = 0. There was no significant effect of freeze position, F(2, 265) = 1.00, p = .370, η^2_G = .007, or interaction with freezing, F(4, 265) = 1.17, p = .324, η^2_G = .017.

Video Quality

Overall video quality. There was a significant effect of freezing on participants' ratings of the overall video quality of the lecture, F(2, 265) = 93.40, p < .001, η^2_G = .413, such that participants in the five-second freezing condition, t(177) = 9.80, p < .001, d = 1.45, and the fifteen-second freezing condition, t(162) = 12.71, p < .001, d = 1.88, reported lower ratings for the overall video quality of the lecture compared to participants in the no freezing condition. Participants in the fifteen-second condition also reported lower ratings of video quality compared to participants in the five-second condition, t(173) = 2.41, p = .017, d = 0.36. There was no significant effect of freeze position, F(2, 265) = 0.32, p = .730, η^2_G = .002, or any interaction with freezing, F(4, 265) = 0.88, p = .478, η^2_G = .013.

Ability to understand the material. There was a significant effect of freezing on participants' reported ability to understand the material, F(2, 265) = 23.91, p < .001, η^2_G = .153, such that participants in the five-second freezing condition, t(173) = 5.84, p < .001, d = 0.86, and the fifteen-second freezing condition, t(181) = 5.84, p < .001, d = 0.86, reported that the video quality negatively affected their ability to understand the material compared to participants in the

no freezing condition. There was no difference between participants in the five-second and fifteen-second freezing conditions, $t(171) = 0.62$, $p = .533$, $d = 0.09$. There was no significant effect of freeze position, $F(2, 265) = 1.50$, $p = .226$, $\eta_G^2 = .011$, or any interaction with freezing, $F(4, 265) = 0.77$, $p = .546$, $\eta_G^2 = .011$.

Ability to pay attention to the material. There was a significant effect of freezing on participants' ability to pay attention to the material, $F(2, 265) = 33.49$, $p < .001$, $\eta_G^2 = .202$, such that participants in the five-second freezing condition, $t(176) = 6.25$, $p < .001$, $d = 0.92$, and the fifteen-second freezing condition, $t(180) = 7.40$, $p < .001$, $d = 1.09$, reported that the video quality negatively affected their ability to pay attention to the material compared to participants in the no freezing condition. There was no difference between participants in the five-second and fifteen-second freezing conditions, $t(178) = 1.58$, $p = .117$, $d = 0.23$. There was no significant effect of freeze position, $F(2, 265) = 0.05$, $p = .954$, $\eta_G^2 < .001$, or any interaction with freezing, $F(4, 265) = 0.47$, $p = .761$, $\eta_G^2 = .007$.

Enjoyment of the material. There was a significant effect of freezing on participants' enjoyment of the material, $F(2, 265) = 46.39$, $p < .001$, $\eta_G^2 = .259$, such that participants in the five-second freezing condition, $t(172) = 7.31$, $p < .001$, $d = 1.08$, and the fifteen-second freezing condition, $t(165) = 8.47$, $p < .001$, $d = 1.25$, reported that the video quality negatively affected their enjoyment of the material compared to participants in the no freezing condition. There was no difference between participants in the five-second and fifteen-second freezing conditions, $t(179) = 1.09$, $p = .278$, $d = 0.16$. There was no significant effect of freeze position, $F(2, 265) = 0.55$, $p = .581$, $\eta_G^2 = .004$, or any interaction with freezing, $F(4, 265) = 0.64$, $p = .634$, $\eta_G^2 = .010$.

Bivariate Correlations

We examined the bivariate relations between the various dependent variables, which are displayed in Table 13. Again, significance values should be interpreted cautiously. As with the previous experiments, the correlations were generally consistent (e.g., test performance was positively correlated with invested effort and with liking and affect, and negatively related to required effort, negative affect and mind wandering).

Table 13

Bivariate correlation table for Experiment 4

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1. Test Performance																		
2. Effort Invested	.23**																	
3. Effort Required	-.18**	.38**																
4. Overall MW	-.32**	-.44**	-.03															
5. Intentional MW	-.01**	-.28**	-.11*	.45**														
6. Unintentional MW	-.34**	-.28**	.05	.77**	-.22**													
7. JOL Before	.54**	.17**	-.33**	-.31**	-.04	-.31**												
8. JOL After	.69**	.19**	-.33**	-.32**	-.09	-.28**	.76**											
9. Liking (Experience)	.15**	.33**	-.02	-.35**	-.30**	-.17**	.25**	.21**										
10. Liking (Content)	.43**	.31**	-.15*	-.25**	-.12*	-.18**	.53**	.53**	.49**									
11. Liking (Visuals)	.02	.27**	.02	-.23**	-.23**	-.08	.18**	.17**	.41**	.32**								
12. Liking (Instructor)	.04	.33**	.06	-.31**	-.31**	-.12*	.20**	.15*	.46**	.30**	.52**							
13. Watch Again	.20**	.32**	-.10	-.40**	-.22**	-.28**	.29**	.33**	.54**	.47**	.55**	.51**						
14. Positive Affect	.19**	.49**	.10	-.34**	-.20**	-.23**	.26**	.26**	.58**	.51**	.46**	.48**	.63**					
15. Negative Affect	-.19**	-.15*	.10	.19**	.14*	.11*	-.20**	-.11*	-.21**	-.14*	-.09	-.12*	-.11*	-.09				
16. Video Quality (VQ)	-.02	.16**	.01	-.13*	-.19**	.00	.08	.08	.40**	.15*	.26**	.25**	.29**	.29**	-.19**			
17. VQ Understanding	.15*	.19**	-.08	-.20**	-.23**	-.06	.14*	.17**	.37**	.18**	.29**	.31**	.38**	.37**	-.23**	.64**		
18. VQ Attention	.08	.22**	-.02	-.20**	-.25**	-.04	.08	.11*	.39**	.20**	.26**	.26**	.33**	.34**	-.19**	.66**	.79**	
19. VQ Enjoyment	-.02	.15*	.00	-.12*	-.24**	.04	.01	.03	.44**	.20**	.33**	.30**	.37**	.33**	-.14*	.69**	.69**	.75**

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

2.4.3 Discussion

Similar to previous experiments, there were no significant effects of freezing or freeze position conditions on test performance, positive affect, and effort. There was no effect of freezing on overall, intentional, and unintentional mind wandering. There was a significant interaction between freezing and freeze position on intentional mind wandering. In the back-loaded condition, participants in the fifteen-second freezing condition reported more intentional mind wandering compared to participants in the five-second freezing and no freezing conditions. With respect to the effect of freeze density, there was a significant effect of freeze density where participants reported more overall mind wandering in the dense periods of freezing compared to the periods of no freezing. There was also a significant interaction between freeze density and freeze duration. The interaction reflected a change in the effect of freeze density across the freeze duration conditions. In the no freezing condition, there were no differences in the proportion of mind wandering between the periods of dense freezing and no freezing. However, in the five-second and fifteen-second freezing conditions, participants reported more overall mind wandering in the dense periods of freezing compared to the periods of no freezing. This pattern was specific to intentional mind wandering where the effect of freeze density was much larger in the fifteen-second freezing condition compared to the five-second freezing condition. These results suggest that dense freezing periods, particularly long periods, may be contributing to participants' intentional disengagement from the lecture. There were no significant effects of freezing on predicted scores before and after test. There were no effects of freezing on liking the content, visuals, and instructor. There was an effect of freezing on liking the general experience of the lecture

and negative affect, such that participants in the five-second and fifteen-second freezing conditions reported lower liking for the general experience of the lecture and reported higher negative affect compared to participants in the no freezing condition. There was no effect of freezing on the likelihood of watching the lecture again in the same format.

Interestingly, there was also an effect of freezing on participants' perceptions of video quality. Compared to participants in the no freezing condition, participants in the five-second and fifteen-second freezing conditions reported lower ratings for overall video quality and reported that the video quality negatively affected their ability to understand the material, their ability to pay attention to the material, and their ability to enjoy the material. These results were also not impacted by freeze position, suggesting that the mere presence of freezing seems to be contributing to participants' negative perceptions of video quality rather than the distribution of freezing events.

Chapter 3: General Discussion

Across four experiments, we examined how freezing impacted learning and the broader learning experience in the context of recorded lectures. The effects of freezing on each dependent variable across experiments are shown in Table 14.

Overall, there were no consistent effects of the freezing condition on learning, load, and metacognition. There were effects of the freezing condition on a few affect variables, specifically in Experiments 3 and 4 that featured longer freeze durations, in which participants in the freezing conditions reported lower liking for the general experience of the lecture. However, their negative experiences did not seem to colour specific aspects of the lecture (e.g., lecture content, perceptions of the instructor). With respect to attention (i.e., mind wandering), Experiments 3 and 4 suggest that dense bouts of freezing led to disengagement. This was most apparent in the fifteen-second freezing condition.

In Experiment 4, there was an effect of the freezing condition on participants' perceptions of video quality. Compared to participants in the no freezing condition, participants in the freezing conditions believed that the video quality negatively affected their understanding of the lecture content, their ability to pay attention to the lecture, and their enjoyment of the lecture. These results suggest that participants were, at least, sensitive to the manipulation of video quality and had beliefs about its potential impact. Although their beliefs in the freezing events negatively affecting their ability to pay attention to the lecture and their enjoyment of the lecture were arguably accurate, there were no overall effects of freezing on our measure of learning.

Table 14

Effect of freezing on dependent variables in Experiment 1, Experiment 2, Experiment 3, and Experiment 4

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
1. Test Performance	ns	ns	ns	ns
2. Effort Invested	ns	ns	ns	ns
3. Effort Required	ns	ns	ns	ns
4. Overall MW	ns	ns	ns	ns
5. Intentional MW	ns	ns	ns	ns
6. Unintentional MW	negative effect	ns	ns	ns
7. JOL Before	effect	ns	ns	ns
8. JOL After	ns	ns	ns	ns
9. Liking	ns	-	-	-
10. Liking (Experience)	-	positive effect	negative effect	negative effect
11. Liking (Content)	-	positive effect	ns	ns
12. Liking (Visuals)	-	ns	ns	ns
13. Liking (Instructor)	-	ns	ns	ns
14. Watch Again	ns	positive effect	ns	ns
15. Positive Affect	ns	ns	ns	ns
16. Negative Affect	ns	ns	ns	negative effect
17. Video Quality (VQ)	-	-	-	negative effect
18. VQ Understanding	-	-	-	negative effect
19. VQ Attention	-	-	-	negative effect
20. VQ Enjoyment	-	-	-	negative effect

Note. ns indicates effect of freezing is not significant. "-" indicates that the effect of freezing did not exist for these variables in each experiment.

3.1 The Effect of Freezing on Learning, Load and Metacognition

Overall, there was no effect of freezing on learning. It is possible that the delays caused by freezing were insufficient to disrupt the construction of a proper mental model of the lecture content, as outlined in the CTML framework, or any cost associated with the delays is balanced by the benefits (e.g., learners may have used the delays to finish processing the learning material). Consistent with this interpretation, there was no effect of freezing on load. That is, there were no differences in invested or required effort ratings between the freezing and no freezing conditions. There was also no effect of freezing on metacognition. When participants were asked to predict their performance, they did so accurately in the sense that they also predicted no effect of freezing on their test performance. Interestingly, when asked directly if the video quality impacted their learning, participants in the freezing condition thought that it did so negatively. It is possible that participants hold pre-existing beliefs about how freezing will affect their learning and that this item indexes that belief. On the other hand, the metacognitive predictions before and after the test (which showed no effect of freezing) might index a representation of the lecture in memory. Provided the latter appears unaffected by freezing, these judgments are also unaffected. Thus, the apparent dissociation between these metacognitive judgments likely represents each relying on different information.

3.2 The Effect of Freezing on Attention

There was an effect of freezing on attention. There were no differences in reported overall mind wandering between participants in the freezing and no freezing conditions. However, with respect to freeze density, we did find that participants were mind wandering during points of dense freezing events. According to Moreno's CATLM framework (2006), affective factors impact learning via cognitive engagement (i.e., attention). Individuals may have experienced the freezing events as disruptive and frustrating, leading to negative affect and disengagement. Consistent with this idea, participants in the freezing conditions liked the general experience of the lecture less and experienced higher negative affect compared to participants in the no freezing condition. In addition, across experiments, mind wandering was positively related with negative affect. When asked directly if the video quality impacted their attention, participants in the freezing condition thought that it did so negatively. In this case, participants' beliefs about the impact of freezing on attention were accurate.

3.3 The Effect of Freezing on Affect

As suggested in the mind wandering results, there was an effect of freezing on affect. While there were no differences in ratings of positive affect, liking the content of the lecture, liking the visuals of the lecture, and liking the instructor of the lecture across the freezing and no freezing conditions, there were differences in ratings of negative affect and liking the general experience of watching the lecture. In Experiment 3, participants in the freezing condition reported liking the general experience of the lecture less than participants in the no freezing condition. In Experiment 4, participants in the freezing conditions reported higher

negative affect and again reported liking the experience of watching the lecture less compared to participants in the no freezing conditions. In addition, when asked directly if the video quality impacted their enjoyment of the lecture, participants in the freezing condition believed that it negatively affected their enjoyment. These results are consistent with past research that found that freezing events have the largest impact on user engagement (Dobrian et al., 2011) and that perceptibility of general video quality issues is a major category of concern (Lange & Costley, 2020).

Interestingly, in Experiment 2, participants in the freezing condition reported liking the general experience and visuals of the lecture more than participants in the no freezing condition. Participants in the freezing condition also reported that they would watch the lecture again in the same format. We did not find this positive effect in Experiment 1 nor did we find this effect in the five-second condition of Experiment 4, thus it might reflect a Type I error. A different possible explanation for these results is that participants in the freezing condition were experiencing some form of cognitive dissonance between the freezing events and their experience of the lecture. According to Leon Festinger's idea of Effort Justification (1962), individuals tend to rate tasks or activities more positively after investing a significant amount of effort in completing them. After investing some effort to understand the video lecture with freezing events, it is possible that participants were trying to justify this invested effort by rating the general experience more positively compared to participants in the no freezing condition. Consistent with this idea, we did find a positive relationship between

invested effort and liking the general experience of the lecture. However, there was no effect of the freezing condition on invested effort.

3.4 Other Effects on the "Learning Experience"

With respect to the distribution or position of freezing events in the lecture, there were limited effects. There were no consistent effects of freeze position on learning, metacognition, attention, or affect. There was an effect on load in Experiment 3 where in a front-loaded freezing condition, participants in the no slides condition were investing less effort into the lecture compared to participants in the slides condition. In general, freeze position does not seem to modulate the effect of freezing.

In Experiment 3, we also examined whether the effect of freezing would be modulated by the presence of lecture slides. While the presence of slides did not appear to moderate the effect of freezing significantly, participants who had slides did better on the test, liked the experience, visuals, and instructor of the lecture more, and experienced higher positive affect compared to those who did not have slides. These results are consistent with research on verbal redundancy in multimedia learning. Verbal redundancy describes the concurrent presentation of the same information through different presentational modes such as spoken narration and written text (Sweller, 2005), which often occurs in video lectures. Most research has demonstrated that students who learned from verbally redundant, spoken-written presentations outperformed those who learned from spoken-only presentations (Adesope & Nesbit, 2012). These findings, along with our own findings, suggest that visual information can have substantial effects on a student's learning experience. From a design

perspective, instructors should consider using visual information to supplement verbal information in their lectures.

3.5 Conclusion

Across four experiments, we found no effect of freezing on learning, load, and metacognition. However, we found that dense bouts of freezing can lead to increases in mind wandering and freezing can negatively impair the lecture experience. In addition, participants appear to believe freezing negatively impacts their learning, attention, and enjoyment. From a practical point of view, freezing should be considered in the context of instructional design. For example, one way to reduce the likelihood of freezing events is to allow students to download lectures to their personal devices. Future work investigating other video quality issues, such as audio-video desynchronization and video resolution, and their impact on the learning relevant variables we explored here (i.e., LLAMA) would be valuable.

References

Adesope, O. O., & Nesbit, J. C. (2012). Verbal redundancy in multimedia learning environments: A meta-analysis. *Journal of Educational Psychology*, *104*(1), 250.

Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education*, 15(4).

Bruning, R. H., Schraw, G. J., & Ronning, R. R. (1999). *Cognitive psychology and instruction*. Prentice-Hall, Inc., One Lake Street, Upper Saddle River, NJ 07458.

Canadian Digital Learning Research Association. (2019). *National Survey of Online and Digital Learning 2019 National Report*.

Canadian Digital Learning Research Association. (2021). *National Survey of Online and Digital Learning 2021 National Report*.

Dobrian, F., Sekar, V., Awan, A., Stoica, I., Joseph, D., Ganjam, A., Zhan, J., & Zhang, H. (2011). Understanding the impact of video quality on user engagement. *ACM SIGCOMM computer communication review*, *41*(4), 362-373.

Festinger, L. (1962). Cognitive Dissonance. *Scientific American*, *207*(4), 93–106. http://www.jstor.org/stable/24936719

Lange, C., & Costley, J. (2020). Improving online video lectures: learning challenges created by media. *International Journal of Educational Technology in Higher Education*, *17*(1), 1-18.

Loderer, K., Pekrun, R., & Lester, J. C. (2020). Beyond cold technology: A systematic review and meta-analysis on emotions in technology-based learning environments. *Learning and instruction*, *70*, 101162.

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

Moreno, R. (2006). Does the modality principle hold for different media? A test of the method-affects-learning hypothesis. *Journal of Computer Assisted Learning*, *22*(3), 149–158.

Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday Attention: Variation in Mind Wandering and Memory in a Lecture. *Applied Cognitive Psychology*, *26*(2), 234–242.

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60.

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

Sweller, J. (2005b). The redundancy principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 159–167). New York, NY: Cambridge University Press.

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*(6), 1063–1070.

Appendix A: Mind Wandering Probe Timings

The timing of each probe for each lecture in each condition.

	Probe 1	Probe 2	Probe 3	Probe 4
Ghanes00 (No Freezing)	1:37	3:55	5:20	9:13
Ghanes00 (Five Seconds, Front-loaded)	2:22	5:35	7:00	10:53
Ghanes00 (Five Seconds, Uniform)	1:52	4:30	6:10	10:43
Ghanes00 (Five Seconds, Back-loaded)	1:37	3:55	5:20	10:15
Ghanes00 (Fifteen Seconds, Front-loaded)	4:18	9:54	11:21	15:16
Ghanes00 (Fifteen Seconds, Uniform)	2:31	6:02	8:21	14:38
Ghanes00 (Fifteen Seconds, Back-loaded)	1:37	3:55	5:20	12:45
Music (No Freezing)	2:27	4:10	5:17	6:52
Music (Five Seconds, Front-loaded)	3:52	5:50	6:57	8:32
Music (Five Seconds, Uniform)	2:52	4:55	6:18	8:12
Music (Five Seconds, Back-loaded)	2:27	4:10	5:17	7:37
Music (Fifteen Seconds, Front-loaded)	6:45	9:14	10:23	11:59
Music (Fifteen Seconds, Uniform)	3:44	6:29	8:22	10:58
Music (Fifteen Seconds, Back-loaded)	2:29	4:13	5:21	9:27

Appendix B: Comprehension Test Questions

List of the 11 test questions in each lecture in the order participants received them. The exact timing for the question's content in the no freezing video is noted first. Question 11 from the Music lecture does not contain a time as it requires participants to retrieve facts from throughout the lecture. "OR" represents the open response questions, "MC" represents the multiple choice questions. Questions with * next to the type of response were considered inference questions.

Ghanes00 Lecture

1) **06:35 – OR*.** In Ghanes00, walking down the roads in towns like S8, NW8, and E8 you would encounter less diversity in terms of fashion (e.g. variations in fabric, dyeing techniques) than towns like S1, NW1, and E1. Why might this have been the case?
2) **03:20 – MC*.** In some of the records found, there was a story of an individual making an important and time-sensitive trip from town N3 to E3. Assuming the traveler traveled along the typical route, which places would they likely pass through?
3) **05:30 – OR*.** How might overhunting of the Nopa45 (e.g., to extinction) influence fashion in Ghanes00? Explain your answer.
4) **09:29 – MC.** Which of the following contributed to the downfall of the Ghanes00 empire?
5) **02:43 – MC*.** Which one of these towns were likely the closest to the Capitol?

6) **04:05 – MC.** Brum62s had a tendency to want to remain close to RiverO99, which was their birthplace.
7) **09:00 – OR.** Why did the people of Ghanes00 use their method of appointing new mayors?
8) **04:35 – MC.** The word Ghanes originally meant
9) **07:50 – OR.** The lecture described the selection process for representatives in the Jemar48. After individuals were selected by passing rigorous written and oral examinations, what were the next steps in the selection event?
10) **00:58 – MC.** Ghanes00 was undiscovered for so long because it was geographically too far to be explored.
11) **05:20 – MC.** Which one of these is NOT a characteristic of the symbiotic relationship between Leni71 and Nopa45?

Music Lecture

1) **02:41 – OR*.** A-6, ? , C-7, D-7. What note should be in place of the question mark?
2) **06:59 – MC.** A combination of two notes is called a(n) ___.
3) **03:21 – MC.** What is the ratio in frequency of E6 to E5?
4) **00:55 – MC.** The higher the frequency, the higher the pitch. True or false?
5) **04:48 – MC*.** Which note has the same pitch as B7 sharp?
6) **06:58 – MC*.** What are the intervals of the following two dyads: E flat to F, D sharp to F?

7) **07:22 – OR.** Why do dyads with certain intervals sound better than others?

8) **07:22 – MC.** Generally, what intervals sound the best?

9) **01:04 – OR.** Why do we give names to frequently used pitches?

10) **04:34 – MC*.** True or False: A3 sharp is the same as B3 flat.

11) **N/A – MC.** Which of the following statements is correct?