MEASUREMENT OF MIND WANDERING IN NATURAL AND UNNATURAL SETTINGS

by

Trish Varao-Sousa

Hons. B.A., University of Waterloo, 2012

M.A., University of British Columbia, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

November 2019

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Measurement of mind wandering in natural and unnatural settings

submitted by Trish Varao-Sousa in partial fulfillment of the requirements

the degree of Doctor of Philosophy

in Psychology

Examining Committee:

Alan Kingstone, Psychology

Supervisor

Rebecca Todd, Psychology

Supervisory Committee Member

Luke Clark, Psychology

University Examiner

Leah Macfadyen, Language & Literacy Education

University Examiner

Additional Supervisory Committee Members:

Steve Heine, Psychology

Supervisory Committee Member

Abstract

Mind wandering (MW) is defined as a lapse of attention where one's thoughts are focused inward, thus drawing attention away from the current task. MW has been reported to occur regularly in tasks ranging from simple visual search-paradigms to real-time driving. MW research has been measured most often within lab settings, however an understanding of how MW occurs in more naturalistic everyday tasks is highly valuable as well. This dissertation was motivated by research on the naturalness of experimental design (Tunnell, 1977) and the practice of cognitive ethology, which aims to understand cognition and behaviour under real-world settings (Kingstone, Smilek, & Eastwood, 2008; Kingstone, Smilek, Ristic, Kelland Friesen, & Eastwood, 2003). The goal was to investigate whether the methods, measurements, and settings used to study mind wandering reflect the natural experience as it occurs inside and outside of the lab. In Experiments 1-5, I examine whether MW report style (self-caught versus probe-caught) influences reports made in terms of frequency or content characteristics. Experiments 6-8 extended this investigation to more naturalistic settings. Experiments 9 and 10 consider the role of distraction in MW, within lab and uncontrolled settings. Results from this body of work reveals that MW is relatively stable regardless of whether it is revealed through probe-caught or self-caught methodologies, that peer presence but not lecture features influence MW rates, and that other forms of inattention (i.e., distraction) do not necessarily occur at the same rate as MW. The findings of this thesis open new avenues and methods for the investigation of MW in controlled and natural settings.

Lay Summary

While completing everyday tasks, such as reading or driving, people regularly report lapses in attention. These lapses are generally known as mind wandering (MW). The goals of this thesis are to examine: a) how our understanding of MW varies as a function of how it is measured; b) how mind wandering occurs in a variety of everyday tasks, such as watching lectures and reading; and c) what role distraction plays in MW across lab and natural settings.

Preface

All of the work in the present document was conducted in the Brain and Attention Research lab at the University of British Columbia where Alan Kingstone is the Principal Investigator. I conducted and/or supervised all data collection. I analyzed the data and was the primary writer for all the manuscripts. My supervisor, Alan Kingstone, assisted with project conceptualization and suggested edits to the manuscripts. The experiments reported were approved by the University of British Columbia's Behavioural Research Ethics Board [Toward a More Natural Approach to Attention Research 1-200: H10-00527].

A version of Experiment 5 in Chapter 2 has been published. Varao-Sousa, T. L., & Kingstone, A. (2018). Are mind wandering rates an artifact of the probe-caught method? Using self-caught mind wandering in the classroom to test, and reject, this possibility. *Behavior research methods*, 1-8.

A version of Experiment 6 in Chapter 3 has been published. Varao-Sousa, T. L., & Kingstone, A. (2015). Memory for lectures: How lecture format impacts the learning experience. *PloS one*, *10*(11), e0141587.

A version of Experiment 7 in Chapter 3 has been published. Varao-Sousa, T. L., Mills, C., & Kingstone, A. (2019). Where You Are, Not What You See: The Impact of Learning Environment on Mind Wandering and Material Retention. In *Proceedings of the International Conference on Learning Analytics and Knowledge.* Caitlin Mills assisted with data analyses and revisions of the manuscript.

A version of Experiment 9 in Chapter 4 has been published. Varao-Sousa, T. L., Smilek, D., & Kingstone, A. (2018). In the lab and in the wild: How distraction and mind wandering affect attention and memory. *Cognitive research: principles and implications, 3*(1), 42. Dan Smilek assisted with experiment design and revisions of the manuscript.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I am grateful to my mentors of the past (Dan Smilek, Colin MacLeod, Jon Carriere, Tanya Jonker & Paul Seli) without their encouragement I may never have pursued this crazy journey.

To all the BAR Lab members: THANK YOU for seeing me through this adventure, for encouraging me to endure, and for all the laughs.

Alan, thank you for encouraging me to pursue not only the low-hanging fruit, but also the pie-in-the-sky, ambitious, ecological, and personally meaningful research. Your support, humor, patience, and tolerance of my cat-obsession have made working with you a wonderful learning and growing experience. I am so grateful to have been part of the lab. Guess we're really off to the races now.

And, of course, thank you to London.

This work was supported by the Natural Science and Engineering Research Council of Canada and the University of British Columbia.

Dedication

To everyone who has supported me along the way: thank you for believing in me, when I didn't believe in myself.

Chapter 1: Introduction

It is likely that you can recall an instance of driving on a familiar route when you suddenly realised that you had reached the destination without being sure how you got there. At some point while driving, your focus shifted from the road to other thoughts, such as an event from earlier in the workday, or imagining a holiday. Likewise, while reading an article you may get to the end of a paragraph to realise that your eyes were moving over the words without processing them - your thoughts drifted to other considerations, such as groceries needed for dinner. These scenarios are examples of *mind wandering* (MW), a frequently reported behaviour where one's thoughts shift away from the primary task (i.e., driving or reading in the above examples) to internal thoughts (Smallwood & Schooler, 2006). Though mind wandering is reported to occur in response to an experimental query roughly 30-50% of the time, MW rates can vary significantly based on the complexity of task and setting.

This dissertation was motivated by research on the naturalness of experimental design (Tunnell, 1977) and the practice of cognitive ethology, which aims to understand cognition (e.g., perception, sensation, learning, awareness, etc.) and behaviour under real-world settings (Kingstone et al., 2008, 2003). The primary goal of cognitive ethology is to raise awareness around whether theories and principles investigated in lab-based experiments will generalize to more complex natural settings, and suggests that researchers wanting to generalize findings from either environment work iteratively through contexts of varying naturalness and control to determine generalizability (Kingstone et al., 2008, p. 317). The goal behind my own research is to investigate whether the methods, measurements, and settings used to study MW reflect the natural experience as it occurs outside of the lab.

1.1 Consequences of a wandering mind

One reason that researchers are so interested in studying MW is that the behaviour can impact primary task performance. Increased reports of MW have been found to negatively impact performance on a variety of tasks relating to attention, including driving, learning, vigilance, visual search, and information retention (Carriere, Cheyne, & Smilek, 2008; Franklin, Smallwood, Zedelius, Broadway, & Schooler, 2016; Hollis & Was, 2016; Kam & Handy, 2013; Mrazek et al., 2012a; Ottaviani & Couyoumdjian, 2013; Smallwood, Baracaia, Lowe, & Obonsawin, 2003; Smallwood, McSpadden, & Schooler, 2008; Stawarczyk, Majerus, Maquet, & D'Argembeau, 2011). It is believed that a performance cost occurs because MW involves a shift of attention away from the primary task, meaning the task cannot be completed as efficiently or as error-free as when attention is focused on the task (Smallwood & Schooler, 2006). For students, one familiar example is sitting in a lecture hall, and suddenly realising the information being presented does not make sense due to a passing thought of something unrelated to the lecture. The MW occurrence (thinking about something unrelated) leaves a "gap" in the lecture information acquired, meaning that later testing on the lecture material presented around the point of MW may result in a performance decrement (Lindquist & McLean, 2011; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Szpunar, Moulton, Schaeter, Schacter, & Schaeter, 2013; cf. Wammes, Seli, Cheyne, Boucher, & Smilek, 2016). In another example, MW while driving an automobile increases the likelihood of accidents, via risky driving behaviours such as increased lane deviations and slower brake times (Baldwin et al., 2017; Cowley, 2013; He, Becic, Lee, & McCarley, 2011; Qu et al., 2015; Yanko & Spalek, 2013). The potential for MW to negatively impact any range of task performances demonstrates the value of a careful analysis of this cognitive state.

1.2 Measuring mind wandering

To track the internally-driven cognitive state of MW, researchers have used a range of external experimental techniques, including eye tracking, self-reports, and neuroimaging (brain imaging). The choice of method depends largely on the research question and task being examined. For example, eye tracking while completing a task in the lab is much more feasible than obtaining fMRI recordings in a classroom lecture. The most common method used to sample MW is subjective self-report, which can be either probe-caught or self-caught (Experience Sampling Method: Csikszentmihalyi & Larson, 2014; Csikszentmihalyi, Larson, & Prescott, 1977).

In the probe-caught methodology, MW is measured by interrupting the individual during the task (typically in response to audio or visual probes/prompts) to sample where an individual's attention is focused. The value of this method is that the researcher can i) manipulate probe intervals to measure changes in MW reports over time, ii) insert MW probes at specific time points to examine the impact on task performance, and iii) the proportional measure of MW frequency can be easily compared across tasks. This method has often been used in lab experiments of MW, and work by Kane et al. (2007) provided early evidence that the probe-caught MW methodology can also be used to capture MW in everyday settings. By providing participants with palm pilots that generated MW prompts at random intervals over 7 days, the researchers discovered that in everyday settings individuals indicate MW in response to 30% of prompts. The value of using the probe-caught method in everyday settings is that it allows for real-time recording of thoughts, which "enhanc[es] reliability and ecological validity [by allowing for] … assessment of contextual influences on experience" (Kane et al., 2007, p. 615). In sum, research using the probe-caught methodology allows for versatile, task-independent monitoring of MW both in the lab and in daily life, and has revealed that in individuals are frequently off-task, with MW rates typically ranging between 30-50% (Galéra et al., 2012; Kane et al., 2007a; Lindquist & McLean, 2011; Mooneyham & Schooler, 2013; Risko, Anderson, et al., 2012; Seli, Konishi, Risko, & Smilek, 2018; Thomson, Seli, Besner, & Smilek, 2014; Wammes, Boucher, Seli, Cheyne, & Smilek, 2016; Wammes & Smilek, 2017). While participants in daily life may not have automated prompts questioning their current attentional state, the use of probes may be likened to having an external prompt that forces one to re-evaluate current attentional focus (e.g., a colleague asking you a question while you are writing a report, or a loud noise from an outside space). In this way, the probe-caught method may reflect some instances in which one is “caught” mind wandering by an external source.

The alternative self-reporting option is self-caught sampling, in which participants must first become self-aware that their mind is wandering and then report this lapse of attention. In lab-based experiments the range of self-caught reports varies between 1-40 reports during 30 minute reading tasks (Jackson & Balota, 2012; Reichle, Reineberg, & Schooler, 2010; Sayette, Reichle, & Schooler, 2009; Varao-Sousa, Solman, & Kingstone, 2017). The value of this method is that there is no upper bound on the number of reported occurrences, thus the researcher obtains all possible instances of MW and not just those occurring at the time of a probe; and reporting is more akin to the natural way individuals catch themselves MW in daily life. The high degree of ecological validity (generalizability of research experiments to real life behaviour and settings) afforded by the self-caught method is one of its key advantages (Chisholm, Chapman, Amm, Bischof, Smilek & Kingstone, 2014; Kingstone, Smilek, & Eastwood, 2008; Risko, Richardson & Kingstone, 2016). On the other hand, the lack of experimental control means that researchers cannot manipulate whether MW reports line up with specific critical task events, nor can they control the large variability that comes alongside relying on participant meta-awareness to report instances as they occur. Another major criticism of the self-caught method is that it requires participants to “monitor” their attention continuously, and the level of meta-awareness required to catch and report mind wandering instances has not been investigated. These negative aspects may have discouraged researchers from using the self-caught method.

As a result, relatively little work has been done to investigate how the self-caught method stacks up against the probe-caught method, or whether important information is being lost when opting to use one method over the other. In the few experiments that have used both methods concurrently, participants are sometimes “caught” MW by a probe before they self-catch, suggesting that one may not always have awareness of their MW until they are probed (Jackson & Balota, 2012; Reichle et al., 2010; Sayette et al., 2009; Schooler, Reichle, & Halpern, 2004; Smallwood, Fishman, & Schooler, 2007). While probe-caught sampling is not without flaws (e.g., it interrupts one’s current cognitions), it remains the most common measurement tool for MW, both in and out of the lab.

1.3 Thesis overview

Despite the diversity of MW research, an abundance of the work has been done in highly-controlled lab settings. Furthermore, and as indicated above, researchers tend to favor the probe-caught method, possibly overlooking the value of using the self-caught method. The present thesis tackles these two issues over the course of 10 experiments. By using tasks that are more representative of natural, everyday behaviours (e.g., lecture watching, listening to audiobooks) and experimentally comparing features of self-caught and probe-caught MW reports, this thesis seeks to i) determine whether the self-caught method taps into a unique experience beyond what is measured by probe-caught reports, and ii) thoroughly explore MW within more natural tasks and settings than traditionally used. The next section highlights the importance of considering dimensions of naturalness in experimental design, as supported by research demonstrating that cognitive behaviours are not always equivalent between laboratory and everyday settings.

1.3.1 Cognitive ethology and task naturalness

When studying behaviours inside the lab, researchers can use highly controlled methods to determine the mechanism underlying a behaviour, although this may sacrifice certainty around whether the findings generalize to other paradigms or environments. On the other hand, when studying behaviours outside the lab, researchers use more naturalistic ecologically valid designs and methods, which allow for greater external validity, but may have low internal validity due to the lack of control over specific variables (Barker, 1968; Barkley, 1991; Burgess, Alderman, et al., 1998; Kingstone, Smilek, & Eastwood, 2008; Kingstone, Smilek, Ristic, Kelland Friesen, & Eastwood, 2003; Kuehner et al., 2017; Bronfenbrenner, 1977).

There have been many theoretical and empirical discussions regarding the degree to which results found within a lab setting can be considered ecologically valid. Barker (1968) argued that contextual knowledge fundamentally influences the ability to interpret or predict human behaviour, as behaviour is influenced by what is appropriate in a given setting (e.g., a library versus a sporting event). Barker argued that an ecological approach is needed within the field of psychology, with more research on behaviour in the "real world" instead of only in lab settings. On the other hand, some researchers have argued that ecological validity does not ensure generalizability and that lab-based tasks provide the best foundation for theoretical and generalizable work (Banaji & Crowder, 1991; Banaji, 1989). Research by Mook (1984) suggests that many in-lab experiments are designed without the intention of generalising to real-life situations, and that there is still high value in understanding what does happen within a lab.

Certainly, much valuable research has come from controlled in-lab studies and some theoretical work may not be supported by an ecologically valid design, but the inability to generalize these works to natural situations can be a serious limitation (Kingstone et al. 2008). The goal of the present thesis focuses on understanding how researchers might measure the naturally occurring phenomenon, mind wandering, in everyday settings and tasks. As will be discussed below, controlled lab-based tasks, when tested against real-life situations, often fall short in being ecologically valid (see Fawcett, Risko, & Kingstone, 2015).

In one demonstration of the disconnect that can occur between lab and real-world settings, Anderson and Brown (1984) reported that individuals have a higher heartrate when immersed in an actual versus artificial gambling environment, and that the correlation between sensation seeking and bet size was only present within the real-life setting. Another recent instance of lab-based research failing to scale to real-world situations concerns social attention. Lab-based experiments have demonstrated a reliable, internally-valid tendency for people to look frequently at the eyes of the other, when the human stimuli were depicted in well controlled experiments using photos and videos (for review see Risko, Laidlaw, Freeth, Foulsham, & Kingstone, 2012). However, when observers are confronted with real people, the opposite pattern of behaviour is often observed, with people avoiding looking at the eyes of others (Foulsham, Walker, & Kingstone, 2011; Foulsham, Chapman, Nasiopoulos, & Kingstone, 2014; Foulsham, Cheng, Tracy, Henrich, & Kingstone, 2010; Laidlaw, Foulsham, Kuhn, & Kingstone, 2011). Collectively this research suggests that the use of stimuli that lack real-world meaning may fail to generalize to everyday behaviour.

Interestingly, MW research, at least initially, appeared to be an exception to this type of disconnect between lab-based and real-world behaviour. For instance, Grodsky and Giambra

(1990) reported that MW rates were positively correlated at the within-subject level when different dimensions of task naturalness were examined (i.e., a vigilance task and a reading task). Reinforcing the idea that MW may be stable across experimental settings, Ottaviani and Couyoumdjian (2013) found that in-lab MW rates were positively correlated with real-time experience sampling tracking that occurred one year later. However, in more recent years, research demonstrating discrepancies between MW rates in the lab and everyday life have begun to surface.

In Wammes and Smilek (2017), participants reported MW while viewing a lecture in two settings: live as part of an ongoing enrolled course, or a video-recorded version as part of a lab-experiment. Results indicated that students who viewed the live lecture did not report changes in the degree of MW over time, whereas those who watched the video-recording reported increasing rates of MW as the lecture progressed. In another demonstration of lab-life divergence, participants reported MW while completing in-lab executive control tasks (e.g., working memory, and attention restraint tasks) (Kane et al., 2017). The lab-based MW reports were compared to reports of MW collected during a week in participants' day-to-day lives. Results revealed that some individual difference factors, such as working memory capacity (WMC) and personality, predicted MW rates in the lab, however, the same factors were not predictive of MW in daily-life. From these findings, Kane and colleagues conclude that "mind-wandering theories based solely on lab phenomena may be incomplete" (p. 1271). In a similar design, Unsworth and McMillan (2017) investigated whether cognitive ability (mental capacities such as WMC, fluid intelligence, and attention control measures) could predict MW during daily academic studying, and whether MW was predictive of academic performance. After assessing participants' cognitive ability (WMC, fluid intelligence, and attention control measures), participants were asked to track their MW events while studying over the course of a week. Unsworth and McMillan found that everyday MW rates were unrelated to cognitive ability as measured in the lab, and MW rates were not predictive of academic performance.

These recent reports of discrepancy in MW results found in and out of the lab question the reliability and generalizability of MW reports. One goal of this thesis is to contribute to the MW literature with experiments that use natural everyday tasks (i.e., reading versus simple attention paradigms like the Sustained Attention to Response Task (SART)), and when possible, testing MW within natural everyday settings (i.e., within classrooms). The operationalization of 'naturalness' is informed by Tunnell (1977), who applied it to the setting and treatment (task) of experimental research design. Table 1.1 provides an example of how this 2x2 categorization can be applied to MW research designs. An unnatural task and lab experiment would be one wherein researchers are investigating relationships at the most controlled dimension of experimental design, such that manipulations are less noisy and settings are less likely to be affected by unpredictable external events. These tasks rarely represent a behavior one would complete in daily-life, and thus how MW is influenced by factors in unnatural tasks and settings may not align with outcomes in natural tasks and/or settings. None of the experiments in the present thesis fall into this most controlled category of design. MW reports during naturalistic tasks in the lab may lead to greater generalizability of results as they involve tasks that participants are more familiar with and likely to experience. In the present thesis, 8 of the 10 experiments fall into this category. One constraint of tasks conducted in the lab is the fact that participants are denied the freedom that is available in everyday settings. For example, a lecture viewed within the lab lacks the self-directed nature of online learning, nor is the participant likely to have equivalent motivation or interest in the topic as a registered lecture. Furthermore, simply choosing to use naturalistic tasks may not fully capture natural attention behaviours. For example, many researchers have used reading tasks to examine MW behaviour but they have artificially constrained how the material is presented (e.g., word-by-word reading; Foulsham, Farley, & Kingstone, 2013; Franklin, Smallwood, & Schooler, 2011; Jackson & Balota, 2012; Smallwood, Nind, & O'Connor, 2009). The decision to manipulate the "naturalness" of a natural task may have an impact on task behaviours, as demonstrated by Varao-Sousa et al. (2017) who found that while reading, individuals routinely respond to MW by re-reading text, a behaviour that would not be captured in a word-by-word paradigm. While examining everyday life settings, the dimension of unnatural tasks would include a task taking place outside of the lab but involving an experimenter-controlled manipulation, thus creating an artificial or constrained situation. This level is used for two experiments, and two manipulations in my thesis. Finally, the dimension of natural tasks within everyday settings allows researchers to examine thoughts occurring during unconstrained activities that have real consequences, however this dimension relies on participant compliance of responding to prompts or surveys in their own time. No experiments in my thesis used this dimension of design.

Dimension		**Task**	
		Unnatural	*Natural*
Setting	*Lab (Unnatural)*	Completing the Flanker Task (a measure of attentional selection) in the lab	Reading a novel (natural task) in the lab
	Everyday life (Natural)	Attending a lecture (natural setting), with lecture format manipulated (e.g., live versus video recorded professor)	Driving to work

Table 1.1 Examples of natural and unnatural experimental designs (adapted from Tunnell 1977, p. 427).

1.3.2 Experimental design and research questions

The present thesis is comprised of ten experiments that apply various methods, measurements, and settings for studying MW Specifically, I have three broad questions, parsed into three chapters: 1) How can we best measure MW? 2) Does task setting influence MW rates? 3) How does MW compare to other forms of inattention?

The first set of experiments (Chapter 2) examines self-reported measures of MW (self-caught and probe-caught reporting) and whether these two methods capture similar experiences. I begin by testing whether MW rates recorded via the most commonly used method, probe-caught, is influenced by probe modality. The next four experiments in Chapter 2 contrast self-caught and probe-caught reports, motivated by the lack of systematic review of these methods in the existing literature. Ultimately, the goal of this collection of experiments is to compare and contrast MW under these two methods, with the key aim being to determine if and when one method might be preferable over the other. Together, the results from these five experiments suggest that MW can be captured reliably by either self-caught or probe-caught methods, although there are some differences in MW reports between the methods.

Chapter 3 explores whether manipulating the naturalness of tasks impacts MW rate or task performance. I first examine MW in a natural, leisure task via the probe-caught method. Following this, I move out of the lab to examine mind wandering in natural classroom settings. I next examine rates of MW in a task considered by many to be enjoyable and leisurely, video gaming. The goal behind these experiments is to determine the way in which specific more natural features of a task or setting might influence MW and associated performance measures (e.g., stimulus retention). Results from three experiments suggest that while MW rates do not appear to be influenced by the setting, other factors (such as retention performance, interest, motivation) may be impacted.

In the final experimental chapter, I investigate another form of inattention – distraction – and examine if and how it differs from mind wandering. The two experiments were motivated by the idea that certain features of an environment (specifically, those outside the lab) may regularly "capture" attention by means of distraction, resulting in a form of inattention distinct from MW. Results from these two experiments suggest that MW and distraction rates are not equivalent between lab and everyday life settings, and that distraction should be measured alongside MW when examining rates of inattention.

Finally, Chapter 5 summarizes the findings of the thesis and considers their empirical and theoretical impact. Potential limitations and future directions for research are also discussed.

1.4 A note on methodology and results

In addition to the standard null hypothesis significance testing, Bayesian analyses using the BayesFactor package in R (R Core Team, 2015; Rouder, Morey, Speckman, & Province, 2012) were used. The computed Bayes factor value provides a ratio for the data in support of evidence for the null versus the alternative hypothesis. The subscript 01 indicates the value is in support of the null hypothesis (BF_{01}), whereas the subscript 10 indicates the value is in support of the alternative hypothesis (BF_{10}). The greater the Bayes factor, the stronger the evidence in support of the indicated hypothesis – with a Bayes Factor of 3-20 indicating positive evidence and values between 20-150+ indicating strong evidence (Masson, 2011). For example, a Bayes Factor of 15 in the favor of the alternative hypothesis (BF_{10}) would suggest that the data obtained are 15 times more likely to occur under the alternative distribution than the null distribution.

Unless otherwise specified, an alpha level of .05 was used for all statistical tests. A critical reason for including Bayesian analyses is it allows for conclusions to be drawn when the statistical p-value does not allow one to reject the null-hypothesis of no difference. Specifically, Bayesian analyses allow for one to test for evidence *both* for and against a null hypothesis – whereas traditional null hypothesis significance testing (NHST) only investigates for evidence *against* the null. To foreshadow the present thesis, the inclusion of Bayes factor allowed for many of my results to illustrate that self-caught and probe-caught methodologies are not significantly different from each other in terms of capturing MW events, and that indeed the data are more likely to come from similar distributions than different ones.

For all experiments, I report how sample size was determined, all data exclusions (if any), all manipulations, and all measures in the experiment (Simmons, Nelson, & Simonsohn, 2012). For a number of the experiments, no prior literature existed from which sample size estimates could be obtained (i.e., no prior comparisons of self-caught to probe-caught reports). In these cases, a more liberal estimate of effect size was used (typically .3-.5) as the goal was to detect effects that could be found more readily and thus were more likely to be replicable. In short, the goal was to demonstrate the presence of moderate to large effects, and in doing so, the ability to detect small effects may have been compromised.

Chapter 2: Methods of Measuring Self-reports of Mind wandering

The present chapter is comprised of five experiments, and examines self-reported measures of MW (self-caught and probe-caught reporting) and whether these two methods capture similar experiences. As indicated in Chapter 1, MW is most commonly measured via the probe-caught methodology, where the task is periodically interrupted with a visual or auditory cue and participants are asked to report on the focus of their attention. However, this method does not represent how individuals become aware of mind wandering in daily life and probe features have been shown to bias responses (Seli, Carriere, Levene, & Smilek, 2013; Weinstein, De Lima, & van der Zee, 2018). Berthié et al. (2015) reflect that the probe-caught method involves "interrupting the attention devoted to an ongoing task to focus on their own thoughts. The participant is then in a sort of shifting task, consisting in alternating attentional focus from the ongoing task to the mind wandering reporting task" (p. 165). The less used, but more ecologically valid, self-caught reporting method allows participants to report MW freely throughout a task, enabling attention assessments that are more reflective of those that would occur in everyday settings, and allows for greater variability in report frequency.

This chapter begins by testing whether MW rates recorded via the most commonly used method, probe-caught, is influenced by probe modality (Experiment 1). This test is critical as different experiments in this thesis involve different probe and task modalities. Motivated by the lack of systematic review of self-caught and probe-caught methods in prior literature the next four experiments compare these methods. In Experiments 2-4, I examine the myriad of ways in which MW instances may be characterized, and examine whether these characteristics differ when using self-caught versus the probe-caught method. For example, probe-caught MW reports have been categorized by mood (Poerio, Totterdell, & Miles, 2013; Ruby, Smallwood, Engen, &

Singer, 2013; Smallwood, Fitzgerald, Miles, & Phillips, 2009), intentionality (Forster & Lavie, 2009; Seli, Carriere, & Smilek, 2015; Seli, Risko, Smilek, & Schacter, 2016), and temporal orientation of wandering thoughts (Baird, Smallwood, & Schooler, 2011; McVay, Unsworth, McMillan, & Kane, 2013; Smallwood, Nind, et al., 2009; Smallwood et al., 2011). Yet these characteristics have not been examined for self-caught reports, nor has research been conducted to determine whether the two methods are capturing different types of mind wandering. By carefully examining the qualitative traits of MW instances the field can begin to gain a better understanding of what (and perhaps why) individuals MW, as well as whether the way in which MW is reported influences the likelihood of reporting such an instance.

One crucial reason that self-caught and probe-caught methods have not been compared is because the output is reported in non-comparable scales: with probe-caught MW reports examined as a proportion of probes (or percentage), and self-caught MW reports examined as raw rates. Raw rates of probe-caught MW cannot simply be compared against raw rates of self-caught MW, as the probe-caught rates are capped by the number of prompts a participant is given, thus truncating the true range of possible MW instances. As a result, most researchers default to the more controlled probe-caught method, even though as noted above, the self-caught method offers many benefits. Investigations of how these characteristics manifest in self-caught reports could inform why this method may be beneficial in some situations. For example, if self-caught and probe-caught MW differ in their ability to capture the frequency of certain MW characteristics it would be important for researchers to use the appropriate measurement tool to capture these instances. Thus, one of the most valuable contributions from experiments 2-4 is that they provide a solution to the issue of scale conflict. In experiments 2-4, the key measures of interest are the *sub-types* of MW that occur under each of self-caught and probe-caught methods, and these sub-types of MW are examined as proportions of overall MW. For example, by calculating intentional and unintentional rates as a proportion of overall MW for each of probe-caught and self-caught methods (as in experiment 2), I remove the barrier of comparing rates on two different scales.

The final experiment of Chapter 2 (Experiment 5), extends the comparison of self-caught and probe-caught MW beyond a laboratory setting to a natural classroom environment. More explicitly, I test whether the use of the probe-caught method creates an illusion of stability in rates simply due to its regular use. If true, then probes not only influence the specific moment of query within a task, but how participants view, and in turn perform, during the entire experiment. To test this I measure MW using the ecologically valid self-caught methodology, and consider whether self-caught MW rates change when the probe-based methodology is inserted into the experiment.

2.1 Experiment 1: Probe modality

In measuring MW subjectively, the use of probe-sampling is the most widely used method for having participants report their attentional focus. Despite the frequency that this method is used, there has been little consistency in how individuals are probed for their mind wandering, and whether the prompt for attention reports may in fact be impacting their reports of inattention. Within my research I investigate MW across a variety of learning tasks (e.g., silent reading, watching lectures) which vary in the presentation modality, and thus vary in their demands on cognitive processing resources. Baddeley (1992) proposed that "humans have separate channels for processing different information modalities", and research suggests that receiving information in varying modalities can differentially impact performance (Spence &

Driver, 1997; Spence, Nicholls, & Driver, 2001). Specifically, the co-ordination of multiple sensory systems can be constrained by the separate processing systems unable to interact concurrently (Peng, Kirkham, & Mareschal, 2018). One example of this is the Colavita effect wherein visual aspects of a task are more likely to be processed than auditory aspects during an audiovisual task (Colavita, 1974; Sinnett, Spence, & Soto-Faraco, 2007). In the case of MW research, receiving a visual probe (e.g., a text message on a computer screen) while completing an auditory task (e.g., listening to an audiobook) would lead to a mismatch in modality processing and may impact task performance.

Within MW research, the use of audio versus visual probes are often selected for different tasks with no obvious underlining rationale. If it is the case that concurrently processing information of different modalities influences performance, one might expect that the inconsistency between task and probe may pose a problem for the results generated. In three experiments I test whether MW rates vary based on probe-task congruency, and whether the probe-task relationship affects the learning of information near in time to the probe onset. To test the impact of probe style on performance I manipulate whether the probe presentation (visual or auditory) is congruent or incongruent with the task (silent reading or audiobook listening). For example, the presentation of a visual probe during an auditory task constitutes an incongruent scenario. Based on past research indicating that incongruency negatively impacts performance (Spence & Driver, 1997; Spence, Nicholls, & Driver, 2001), I hypothesize that when probe modality does not match task modality, MW rates will be higher, response times will be slower, and performance on tests of material retention will be poorer.

2.1.1 Method

2.1.1.1 Participants

For each of these experiments, G*Power was used to determine the sample size necessary with power set to .80 and alpha set to .05. The goal was to detect an effect of at least d = .50, as this was the minimum effect size that I was interested in examining. For all three experiments, which utilized within-subjects designs, this yielded a required sample size of 34 participants per experiment. All participants were UBC undergraduates who participated in return for course credit (Experiment 1a: 36, Experiment 1b: 37, Experiment 1c: 43). All provided informed written consent before participating and were fully debriefed upon experiment completion. Twelve participants were removed from the analyses (Experiment 1a: 2, Experiment 1b: 4, Experiment 1c: 6) due to failure to properly follow task instructions (e.g., skipping pages) or because no MW was reported.

2.1.1.2 Reading and listening materials

In Experiment 1a, participants read two non-fiction passages, one on the atmosphere (14 pages) and one on fossils (14 pages) (Bryson, 2005). In Experiments 1b (silent reading) and 1c (audio track: 20 minutes, 17 seconds), only the passage on fossils was used.

2.1.1.3 Mind wandering probes

For Experiment 1a, MW probes were presented visually (see Figure 2.1) in one of the reading tasks and auditorily in the other (probe order counterbalanced across participants). Experiment 1b and 1c presented both auditory and visual probes intermixed within the task (7 of each). Visual probes did not include text and instead the visual probe was only the blue box shape (as per Figure 2.1 without text), which flashed twice on the screen (matched to the tone of the audio probe which had two beats).

Figure 2.1. Probe text for Experiment 1a, visual condition.

2.1.1.4 Retention test

Participants in all experiments completed 14 True-False test items related to the material at the end of the task. In Experiment 1a, the tests occurred in between reading blocks.

2.1.1.5 Procedure

The procedure for all three experiments was identical, with the exception of the task that participants were assigned to complete (reading in experiments 1a and 1b, listening in study 1c). Once within the lab testing room, the participant was provided the following definition of MW: "Mind wandering is any thought experienced that is not related to the material being presented", participants were asked to respond via keypress to the probes, 'y' for yes (to indicate MW) or 'n' for no (to indicate not MW). Participants were presented with a sample of the audio tone that would be presented for the audio probe. In the reading experiments, participants were asked to read the passages silently. After completing the reading/listening task participants completed the retention test and were then debriefed on the nature of the experiment.

2.1.2 Results

Table 2.1 includes a summary of the descriptive statistics. The results of the three experiments are discussed together, as the results were highly similar across experiment.

Measure	Experiment 1a (*n* = 34)		Experiment 1b (*n* = 33)		Experiment 1c (*n* = 37)	
	Congruent (V)	Incongruent (A)	Congruent (V)	Incongruent (A)	Congruent (A)	Incongruent (V)
MW reports (%)	37.43 (20.55)	43.26 (25.93)	35.07 (25.56)	39.39 (23.36)	50.19 (19.81)	45.94 (25.60)
Retention test scores (%)	68.68 (14.91)	68.86 (14.76)	75.87 (27.70)	75.06 (28.98)	56.81 (28.32)	66.78 (27.29)
MW RT	2.76 (1.08)	2.45(.40)	2.63 (.97)	2.87 (.86)	2.56 (.56)	2.29 (.55)

Table 2.1. Mean values for each experiment, with standard deviations in parentheses.

2.1.2.1 Mind wandering reports

Across all three experiments, reported MW rates did not differ significantly between congruent and incongruent probe styles (Experiment 1a: $t(33) = 1.67$, $p = .10$, 95% CIs [-.012, .13], BF_{01}: 2.03, $d = .29$, Experiment 1b: $t(32) = .85$, $p = .40$, 95% CIs [-.06, .15], BF_{01}: 5.24, $d = .15$, Experiment 1c: $t(36) = .96$, $p = .34$, 95% CIs [-.05, .13], BF_{01}: 4.99, $d = .16$). Furthermore, an overall test, across experiments, of whether MW rate was impacted by probe type, regardless of task congruency, revealed no significant effect of probe type, $F(1, 101) = 3.59$, $p = .06$, BF_{01}: 2.17, $\eta_p^2 = .03$.

2.1.2.2 Retention test performance

Retention test performance did not differ significantly between congruent and incongruent probe styles for any of the three experiments (Experiment 1a: $t(33) = .05$, $p = .96$, 95% CIs [-.07, .07], BF_{01}: 7.50, $d = .009$, Experiment 1b: $t(32) = .17$, $p = .87$, 95% CIs [-.11, .08], BF_{01}: 7.30, $d = .03$, Experiment 1c: $t(36) = 1.36$, $p = .18$, 95% CIs [-.25, .05], BF_{01}: 3.23, $d = .23$).

2.1.2.3 MW response times (RT)

Response times in Experiments 1b and 1c were influenced by probe modality, where visual probes were responded to more quickly than audio probes (Experiment 1b: $t(32) = 2.99$, $p = 005$, 95% CIs [.08, .39], BF_{01}: .16, $d = .52$, Experiment 1c: $t(36) = 6.45$, $p < .001$, 95% CIs [.18, .35], BF_{01}: 0, $d = 1.08$). Interestingly, for Experiment 1c, this indicates that participants were *faster* when the probe was incongruent to the task. In Experiment 1a, there was no difference in response times between probe format ($t(33) = 1.83$, $p = .08$, 95% CIs [-.07, .04], BF_{01}: 1.59, $d = .32$).

2.1.3 Discussion

Given that MW experiments often use probes as a method for capturing MW, determining whether probe congruency influences reports is relevant when using stimuli in varying domains. The predicted outcome was that when probes come in an opposite modality from the task, there may be a negative effect in switching from one modality processing stream to the other, thus increasing MW rates, increasing reaction time, and impairing performance on recently processed information. However, the three experiments conducted provide no evidence that probe-congruency influences MW reports. As a result of this finding, the remaining experiments in my thesis used a probe modality that was best suited for the task design. The finding that probe-to-task mode congruency does not impact MW outcomes should also simplify experimental design for future probe-caught MW research.

2.2 Experiment 2: MW intentionality

The remaining four experiments in Chapter 2 examine whether report method (self-caught versus probe-caught) influences MW outcomes. To begin this investigation, I examined the role of intentionality on MW reports. Researchers have experimentally examined the frequency of intentional and unintentional MW rates, with most researchers finding that unintentional MW accounts for roughly 65% of reports (Forster & Lavie, 2009; Hines & Shaw, 2013; Seli, Cheyne, Xu, Purdon, & Smilek, 2015; Seli, Risko, Smilek, & Schacter, 2016; Seli, Wammes, Risko, & Smilek, 2015). Moreover, the rate at which unintentional MW occurs may be influenced by task load, with more unintentional MW being reported in difficult as compared to easy tasks (Seli, Risko, & Smilek, 2016), and yet MW intentionality appearing to be uninfluenced by perceptual manipulations (Forster & Lavie, 2009). The prior work on MW intentionality provides a theoretical framework for examining whether intentional and unintentional MW rates vary in more naturalistic tasks. Given the regularity in which MW occurs in educational settings (such as reading or attending lectures), the present experiment investigated whether MW intentionality would differ between reading and viewing tasks, tasks that could be considered high- and low-load, respectively. A second aim of this experiment was to examine whether intentionality would be influenced by the way in which MW was reported (self-caught versus probe-caught). Although researchers have advocated for the importance of distinguishing between MW subtypes in order to understand the origin and impact of MW (Seli, Risko, & Smilek, 2016; Smallwood & Schooler, 2015), research has been exclusively conducted with paradigms using probe-caught reporting.

In the present experiment, I investigate whether rates of intentional and unintentional MW differ when manipulating: 1) the method of reporting (probe-caught or self-caught) between-subject and 2) the task (reading versus video lecture watching) within-subject.

Research Questions:

1) Will overall rates of MW differ between reading and watching tasks?

2) Do rates of MW intentionality vary based on self-report method?

3) Do rates of MW intentionality differ based on the task being completed?

In regard to question 1, I hypothesize that there will be more overall MW (regardless of report method) when watching than silent reading. Silent reading requires high cognitive focus and processing, as compared to passive watching, and so it could be considered to have higher cognitive load – which is often related to reduced MW rates. This is further supported by prior research indicating that reading requires deeper engagement as compared to a more passive tasks (such as lecture watching) (Forster & Lavie, 2009; Smallwood, Fishman, et al., 2007; Varao Sousa, Carriere, & Smilek, 2013). Smallwood, Fishman, et al. (2007) report that MW is more frequent in tasks requiring ‘‘superficial engagement”, and the load hypothesis of Forster and Lavie (2009) suggests that tasks with high load can reduce one’s ability to engage in MW.

Regarding research question 2, within the probe-caught literature there are robust findings that MW is most often reported as unintentional (Forster & Lavie, 2009; Seli, Risko, & Smilek, 2016). While there are no theoretical grounds to speculate that self-caught reports would show a different pattern, if one intentionality type of MW were “easier” to notice (i.e., required less meta-awareness) it would be reported more often during the self-caught session, whereas the probe caught method should capture each type equally. For example, active disengagement from a task (intentional MW) could be more obvious and thus reported more frequently when self-catching as compared to probe-catching. As there has not been any research conducted on intentionality with the self-caught method, this hypothesis is purely speculative.

Regarding research question 3, it is possible that the higher rates of overall MW while watching a video will be driven by increased unintentional MW reports. This outcome would be supported by Forster and Lavie (2009) who found greater rates of unintentional probe-caught MW under low versus high perceptual load. In contrast to this, Seli, Risko, and Smilek (2016) found that participants reported more unintentional MW in a difficult version of the sustained attention to response task (SART), as compared to an easy version. Given this conflict in outcomes regarding intentionality and task load, and that I am examining the rates in a more naturalistic task than Forster or Seli, it is unclear which pattern of results will generalize.

2.2.1 Method

2.2.1.1 Participants

A power analysis for a between-within ANOVA was conducted using G*power (Faul, Erdfelder, Lang, & Buchner, 2007). A predicted medium effect size (d = .25: based on effect sizes reported for intentional/unintentional comparisons across task load in Forster & Lavie, 2009), moderate power-to-detect effects (.80), and a set alpha of .05, indicated that 62 participants would be required. An additional week of data collection was conducted to ensure that these criteria would be met after any unanticipated data exclusions. Eighty-three UBC undergraduate students participated in return for course credit. All provided informed written consent before participating and were fully debriefed upon experiment completion. Nine participants were removed from the analyses (Probe-caught condition: 4 participants; Self-caught condition: 5 participants) due to failure to properly follow task instructions (e.g., skipping pages). The final sample, after exclusions, consisted of 54 women and 20 men (M_{age}= 20.11, SD_{age} = 1.95). Participants were pre-screened to select only those with normal or corrected-to-normal hearing and vision.

2.2.1.2 Reading and lecture materials

Participants read a passage and watched a video-recorded lecture, each for 18 minutes. The passage was one of two excerpts (related to the atmosphere, or fossils) from *A Short History of Nearly Everything* (Bryson, 2003). Reading was self-paced, with one page presented at a time and participants advancing the page via keypress. The lecture was one of two selections from OpenYale online courses (Population Decline in Europe, or Evolutionary Medicine), participants could not pause or rewind the video. Task order and topic were randomized across participants. An important note regarding the selection of task materials is that an executive decision was made to not equate task content. Although using the lecture script for the reading task would have equated the material content across task types and thus created a more controlled experiment, the material was not designed to be read and would therefore represent a less natural experience for the participant. An alternative task would have been to show participants a recording of a book reading session, however this task also felt like an unnatural option which very few individuals would engage in as compared to viewing an online lecture.

2.2.1.3 MW report method

Report method was manipulated between-subject. For participants in the probe-caught condition, an auditory probe sounded 12 times across each task, at a pre-determined pseudo-random time, with probes spaced roughly 60-100 seconds apart (Seli et al., 2013). Participants were required to indicate, via labelled keypress, whether they were: on-task, MW unintentionally, or MW intentionally. For participants in the self-caught condition, participants were asked to indicate, via a keypress, anytime they noticed that they were MW unintentionally or MW intentionally.

2.2.1.4 Retention test

Material retention was assessed via a computerized multiple-choice quiz following each task. Each test had 12 questions associated with the material, with 4 possible answer choices to each question.

2.2.1.5 Ratings

Following each task, participants reported how motivated they were to perform well on the task and how interested they were in the material presented. Participants responded via keypress to a 6-point Likert scale for motivation where 1 represented low motivation and 6 represented high motivation (based on Unsworth & McMillan, 2013), and a 5-point Likert scale for interest where 1 represented low interest and 5 represented high interest (modified from Giambra & Grodsky, 1989).

2.2.1.6 Procedure

After providing informed consent, participants were brought into a testing room where age and gender information was collected. Participants were provided the following definition of MW: “Mind wandering is any thought unrelated to what you are [watching/reading]”. The following definitions were provided for intentional and unintentional MW: “Intentional mind wandering is deliberately choosing to disengage from the task. While you are listening or watching, you decide to start thinking about something else. So in other words, you intentionally began to think about another topic” and “Unintentional mind wandering is spontaneously disengaging from the task. While you are listening or watching, you suddenly notice that your mind is elsewhere, you did not mean to have your attention wander, but it did. So in other words, you were unintentionally thinking about another topic”. Participants were assigned to either the self-caught or probe-caught condition, and completed the video or reading task first, based on a pre-determined randomization. Instructions on responding to probes or responding to self-caught instances were provided based on condition. After each task, participants completed the retention test, and rated their interest and motivation in the material. Once completed, the research assistant set up the next block (reading or watching, based on counterbalance). Upon completion of both tasks, participants were provided feedback on the experiment.

2.2.2 Results

A 2x2 between-within subject ANOVA was conducted on the data, except for the overall mind wandering reports as these were reported on different scales. Table 2.2 provides descriptive statistics for the dependent variables.

Measure	Reading	Lecture Watching
Mind wandering		
Probe-caught proportion (Range 0-1)	.34 (.25)	.48 (.25)
Self-caught count (Range 1-30)	13.92 (8.89)	15.76 (13.1)
Retention performance (Percentage)		
Probe-caught (Range 9-100)	57.8 (17)	58.5 (20)
Self-caught (Range 17-100)	59.4 (19.2)	61.7 (17.7)
Motivation Rating		
Probe-caught (Range 1-6)	3.30 (1.33)	3.24 (1.28)
Self-caught (Range 1-6)	3.21 (1.29)	3.49 (1.26)
Interest Rating		
Probe-caught (Range 1-4)	2.41 (.98)	2.70 (1.10)
Self-caught (Range 1-4)	2.49 (.96)	3.05 (.88)

Table 2.2. Mean reports of dependent measures, with standard deviations in parentheses.

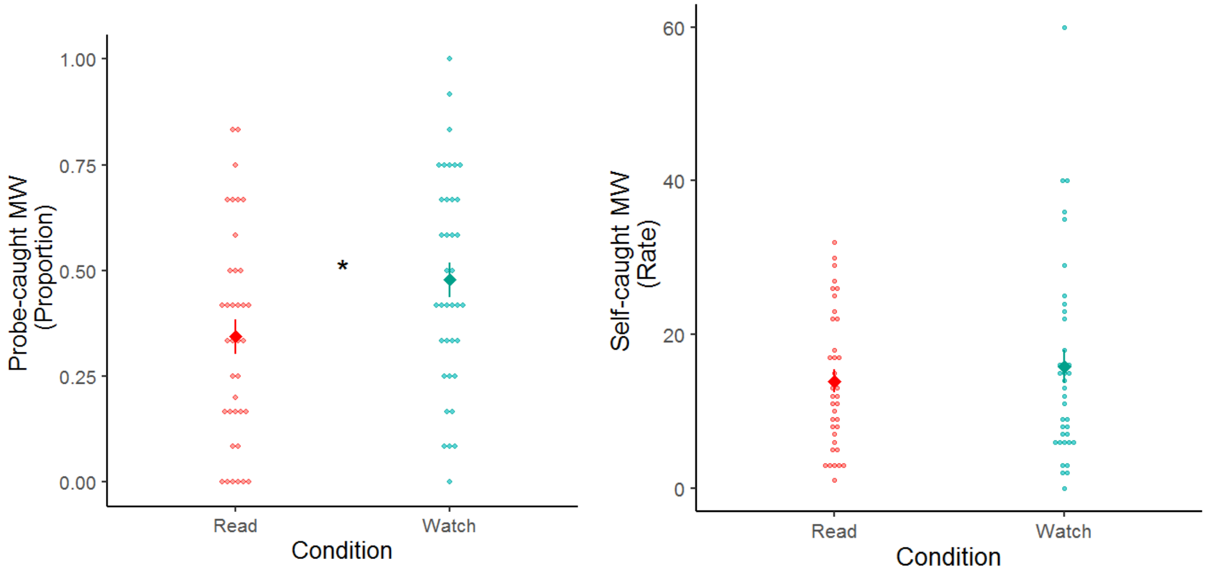


Figure 2.2. MW reports across report method and task type. Bars represent one standard error of the mean.

2.2.2.1 Overall within-subjects MW reports

Overall rates of MW (sum of intentional and unintentional MW) were examined in the original reported scale (proportion for PC, and rate for SC), see Figure 2.2. As hypothesized, participants reported more probe-caught MW when watching a lecture than when reading, $t(36) = 3.19$, $p = .003$, 95% CIs [-.22, -.04], BF_{10}: 11.11, $d = .53$. Self-caught MW reports were not significantly different between the two tasks, although there was a numerical difference in the same direction as probe-caught reports, $t(36) = 1.09$, $p = .29$, 95% CIs [-5.27, 1.59], BF_{01}: 4.43, $d = .18$.

2.2.2.2 Intentional and unintentional MW reports

To circumvent the issue of MW being reported on different scales, analyses were conducted on the proportion of responses reported as unintentional or intentional, as a function of overall mind wandering. I first identified whether participants were more likely to report MW as intentional and unintentional. Consistent with prior research, participants reported significantly more unintentional MW compared to intentional MW, $F(1, 138) = 123.61$, $p < .001$, $BF_{10} > 150$, $\eta p^2 = .66$, regardless of task type or MW report method.

Next, I examined whether rates of unintentionality differed based on task or report method[1], rates are presented in Figure 2.3. There was a main effect of task type, with participants reporting more unintentional MW when reading than watching, $F(1, 64) = 4.34$, $p = .041$, BF_{01}: .78, $\eta p^2 = .13$. There was no effect of report style (probe-caught versus self-caught) on the results $F(1, 64) = .15$, $p = .70$, BF_{01}: 3.82, $\eta p^2 = .13$, nor was there an interaction between the task type and report style, $F(1, 64) = .63$, $p = .43$, BF_{01}: 9.72, $\eta p^2 = .13$.

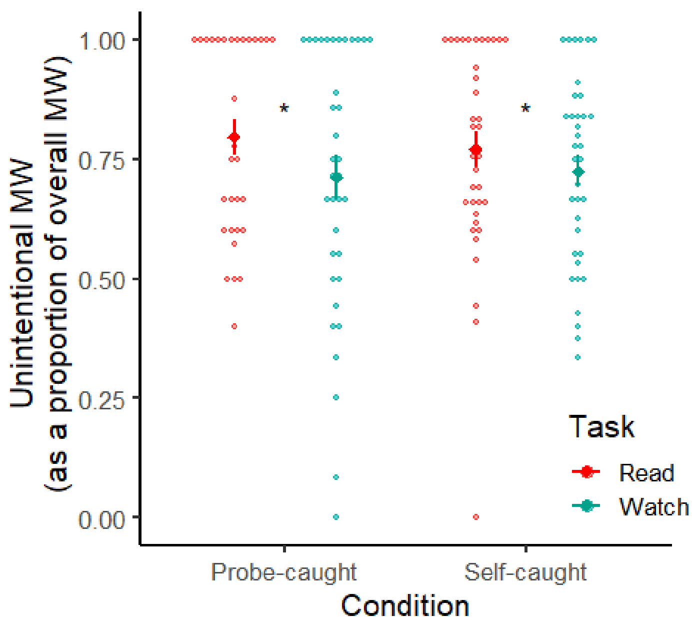


Figure 2.3. Unintentional mind wandering rates, presented as a function of task type and report type. Bars represent one standard error of the mean.

---

[1] Intentional rates could have been examined as well, however because intentional and unintentional rates are ipsative, any results found for one would be identical – but in the opposite direction – for the other.

2.2.2.3 Retention test performance

Retention performance did not differ based on task type $F(1,72) = .61$, $p = .44$, BF_{01}: 5.26, $\eta p^2 = .008$), report method $F(1,72) = .25$, $p = .62$, BF_{01}: 4.17, $\eta p^2 = .004$, nor was there an interaction $F(1,72) = .07$, $p = .80$, BF_{01}: 88.46, $\eta p^2 < .001$.

2.2.2.4 Motivation and interest ratings

Motivation ratings did not differ based on task type ($F(1,72) = .57$, $p = .45$, BF_{01}: 4.22, $\eta p^2 = .008$), report method ($F(1,72) = .09$, $p = .76$, BF_{01}: 3.26, $\eta p^2 = .001$), nor was there an interaction ($F(1,72) = 1.28$, $p = .26$, BF_{01}: 34.35, $\eta p^2 = .018$). Participants reported that the lecture video was more interesting than the reading passage, $F(1,72) = 9.16$, $p = .003$, BF_{01}: .08 (BF_{10}: 12.43), $\eta p^2 = .12$. Interest ratings were not influenced by MW report method ($F(1,72) = 1.46$, $p = .23$, BF_{01}: 2.56, $\eta p^2 = .02$), nor was there an interaction ($F(1,72) = .89$, $p = .35$, BF_{01}: .54, $\eta p^2 = .01$).

2.2.3 Discussion

The present experiment extends past research on MW intentionality to naturalistic tasks and provides information on how patterns of intentionality manifest when using self-caught methods. The goals of this experiment were to answer the following questions: 1) Will overall rates of MW differ between reading and watching tasks? 2) Do rates of MW intentionality vary based on self-report method? 3) Do rates of MW intentionality differ based on the task being completed?

The results of research question 1 were in line with the hypothesis that there would be more overall MW when watching a lecture than silent reading, but only for probe-caught MW. Self-caught MW reports demonstrated this pattern of results, but the outcome was not significant. One possible reason why this finding did not extend to self-caught MW could be that participants did not have enough meta-awareness toward lapses of attention when watching the video lecture. Participants reported that the video lecture was more interesting than the reading material, and perhaps this interest in content reduced the self-awareness. The potential for individual differences in self/meta-awareness to influence reporting outcomes is an important area for future examination.

The outcomes for research questions 2 and 3 were examined within a mixed methods ANOVA. Consistent with prior probe-caught MW research, MW is much more likely to be reported as unintentional (roughly 75% of instances) compared to intentional MW, regardless of task type or MW report method. This outcome indicates that past findings generalize to naturalistic materials, and to the self-caught methodology – further suggesting that one intentionality type of MW is not "easier" to notice (i.e., requiring less meta-awareness) than the other. Results for research question 3 indicated that while overall there was more unintentional than intentional MW, there was more unintentional MW reported when reading than watching. This outcome is in line with the results of Seli et al. (2016), where participants reported more unintentional MW during a high demand versus low demand sustained attention task, but not with Forster and Lavie (2009) who found greater rates of unintentional probe-caught MW under low versus high perceptual load. It is possible that Seli and Forster had conflicting results as difficulty and perceptual load may tap into different resources, thus in the present experiment reading may be a more difficult task than lecture viewing, but perhaps not higher in perceptual load. Specifically, Forster and Lavie (2009) suggest that high perceptual load tasks engage all available attentional capacity whereas during low perceptual load tasks, "leftover" attentional capacity remains and may be captured by MW instances. In their experiment, perceptual load was manipulated by changing the task stimuli to be identical (low load) versus non-identical (high load) distractor items. In contrast to this, Seli et al (2016) manipulated task difficulty by creating an easy, predictable sustained attention task (sequential number order), however the reasoning for this manipulation was not explicitly distinguished from a perceptual load manipulation. Unlike both of these prior experiments, the current tasks were not directly manipulated for difficulty or perceptual load, and the inclusion of a specific measure of task difficulty and/or load would have helped in determining how the reading and watching tasks used in the present experiment differed.

This experiment first replicates the outcome that participants tend to report more unintentional than intentional probe-caught MW, and uniquely demonstrates that this pattern extends to both self-caught MW, and naturalistic tasks. In contrast to predictions, intentionality was not influenced by report method, i.e., the self-catch method does not appear to capture certain MW types any more than probe-caught methods. The present experiment also provides evidence that reading text leads to more unintentional MW than watching lectures, demonstrating that task specific features may influence MW intentionality.

One possible limitation to the experiment is that I did not explicitly measure subjective task demand. That is to say, it is unclear how the difference in cognitive demand between reading and lecture watching was perceived by participants. Future work could investigate what makes a task more (or less) “demanding” (e.g., high in load) and why task load leads to changes in intentional mind wandering. One possibility is that a task that is less demanding makes it easier for participants to “check out” due to the reduction in cognitive resources required to complete the task. The present experiment did not have a sufficiently large sample size to examine the relationship between MW reports and retention performance, thus, unlike Seli et al. (2016) we could not use task performance as a direct measure of difficulty.

The experience of MW is believed to be more nuanced than a simple ‘presence’ or ‘absence’ of attention (Christoff, Gordon, Smallwood, Smith, & Schooler, 2009; Schad, Nuthmann, & Engbert, 2012; Smallwood, McSpadden, & Schooler, 2007; Smallwood & Schooler, 2006). While the present experiment suggests that task features may influence MW reports, intentionality of MW may not be the only way to distinguish different mind wandering instances. Thus Experiment 3 was designed to examine whether probe-caught and self-caught MW rates differ based on the perceived “depth” of the experience. The issue of depth is discussed in detail below.

2.3 Experiment 3: MW depth

The results of the prior experiment suggest MW intentionality is not influenced by MW report method and, most importantly, that individuals do have the meta-awareness to regularly detect and report their MW instances. To provide further evidence that self-caught reporting is a valid method for capturing MW, and to determine whether self-caught and probe-caught instances tap into fundamentally different experiences the following experiment related to MW depth was designed.

Apart from intentionality, MW may differ on features such as thought content or depth /engagement with the thought. Research by van Vugt and Broers (2016) suggests that difficulty in thought-disengagement (“stickiness”) is associated with more frequent MW reports, as individuals will struggle to keep their attention focused on the central task. Related to difficulty in keeping attention focused entirely on the task at hand, Schad, Nuthmann, and Engbert (2012) propose that MW occurs outside of the commonly used yes/no dichotomy, but on a spectrum, with “graded degrees at different hierarchical levels of cognitive processing” (p. 190). Indeed, some researchers have considered that MW may occur on a continuum, and made use of graded scales to measure the "degree" of on-task focus (Dixon & Li, 2013; Farley, Risko, & Kingstone, 2013; Franklin et al., 2011; Seli et al., 2014). Seli et al. (2014) report that MW is associated with increased fidgeting, but specifically that being "completely off task" yielded more fidgeting than other levels of MW. The authors label this most encompassing type of MW as "deep" MW and suggest that different levels of MW may differentially influence performance outcomes. However, the possibility that individuals categorize MW as being either shallow or deep has yet to be empirically examined.

In this experiment, a within-subject design was used to explore whether self-caught and probe-caught MW differ in terms of perceived depth. As in Experiment 2, by examining the proportions of the subtypes of MW (as a function of all MW reported) self-caught and probe-caught measures can be directly compared.

The rate of self-caught MW reports may be influenced by individual levels of mindfulness (sustained attention ability). Mindfulness research would support the hypothesis that shallow MW would more readily be self-caught as the individual is not too deeply "lost" in their thoughts and has higher ability to regulate their attention to the task (Bishop et al., 2006; Mrazek, Smallwood, & Schooler, 2012; Shapiro, Carlson, Astin, & Freedman, 2006). Another possibility is that "shallow" instances of MW may be self-caught less frequently than probe-caught, as one may be less likely to spontaneously detect and report shallow lapses of inattention. If a probe can *always* catch the individual, regardless of whether the instance is shallow or deep, one may be more likely to categorize such an event as MW and thus probes would be equally likely to capture each instance. If this is the case, I would predict fewer shallow self-caught instances relative to probe-caught, as MW may go undetected more frequently when it is shallow due to the fleeting nature of the thought, and the ability to continue with the task without a jarring intrusion. On the other hand, during an instance of deep MW there may be more of the task content missed and thus external indicators that one was off task, making the reporting of deep self-caught MW more common than shallow self-caught MW. Thus, *if* shallow and deep MW occur at an equal frequency, they would be probe-caught an equivalent amount of the time (50% each), however if one is "easier" to detect (i.e., deep MW) then it would be reported more frequently during self-caught sessions.

2.3.1 Method

2.3.1.1 Participants

Sample size was calculated using the following parameters in G*Power (Faul et al., 2007): a within-subjects two-tailed test, with alpha set to .05, power .8, and effect size estimated at .35 (based on differences between intentional and unintentional MW rates in Forster and Lavie, 2009, which could possibly be linked to depth rates). This indicated a required sample size of 67, with data collected for an additional week to ensure a sufficient sample was obtained after exclusions. Seventy-six University of British Columbia undergraduate students participated in return for course credit, however 9 were removed from the analyses due to failure to properly follow task instructions (e.g., missing probes). After exclusions, the sample included consisted of 57 women and 10 men (M_{age}= 21.43, SD_{age} = 4.42). Participants were pre-screened to include only individuals with normal or corrected-to-normal hearing and vision.

2.3.1.2 Materials

The same two pre-recorded lecture videos from Experiment 2 were used, with order of viewing the videos counterbalanced across participants.

2.3.1.3 MW report method

Report method was manipulated within-subject, with order of report method counterbalanced across subject. For the probe-caught block, fifteen audio probes were used in the experiment, with the amount of time between probes ranging from 55-185 seconds. Participants were required to indicate, via a single labelled keypress, whether they were: shallow MW, deep MW, or not MW. If participants stated that they were MW, they were then asked to indicate whether the instance was shallow or deep, via keypress. For the self-caught block, participants were asked to indicate, via a single keypress, anytime they became aware that they were experiencing a deep or shallow MW episode.

2.3.1.4 Retention test, motivation and interest ratings

Content retention, student motivation, and interest were assessed using the same measures as Experiment 2.

2.3.1.5 Perception of MW depth

After completing both sessions, participants were asked to indicate how they defined an instance of MW as deep or shallow. They were also asked to indicate whether the way that they defined these instances differed between the self-caught and probe-caught blocks. Due to the within-subject nature of the task, I hope to take advantage of participants using the same criteria for each of the MW report types, however I needed to confirm that this was indeed the case.

2.3.1.6 Procedure

At the start of the experiment participants provided informed consent before being brought into the testing room. The following definition of MW was provided: "Mind wandering is any thought unrelated to what you are watching". Participants were not provided a definition of "shallow" or "deep" MW as one of the goals of the experiment was to see how these instances were conceptualized by participants and to determine whether they were dissociable experiences, Participants were assigned to complete either the self-caught or probe-caught condition first, based on a pre-determined randomization. At the end of each video, participants completed the retention test, and indicated their motivation and interest in the video content. Upon completion of both tasks, participants completed the MW depth perceptions component and were then debriefed on the nature of the experiment.

2.3.2 Results

Within-subjects analyses were used to analyze the data. Table 2.3 provides descriptive statistics for the dependent variables.

Measure	**Rate**
Overall Mind wandering	
Probe-caught proportion (Range 0-1)	.50 (.23)
Self-caught count (Range 1-47)	13.21 (9.24)
Retention performance (Percentage)	
Probe-caught (Range 16-94)	61.36 (21.26)
Self-caught (Range 11-94)	62.85 (20.97)
Motivation Rating	
Probe-caught (Range 1-6)	3.18 (1.36)
Self-caught (Range 1-6)	3.37 (1.39)
Interest Rating	
Probe-caught (Range 1-5)	2.70 (1.16)
Self-caught (Range 1-5)	2.67 (1.06)

Table 2.3. Mean reports of dependent measures, with standard deviations in parentheses.

2.3.2.1 Deep and shallow MW reports

As in Experiment 2, to circumvent the fact that MW was reported on different scales between self-caught and probe-caught sessions, analyses were conducted on the proportion of responses reported as deep or shallow, as a function of overall MW. I first examined whether shallow and deep MW proportions differed based on reporting type. Participants reported significantly more shallow MW (77.53%) as compared to deep MW (22.47%), $F(3, 183) = 100.24$, $p < .001$, $BF_{10} > 150$, $\eta p^2 = .62$, regardless of report style (probe-caught versus self-caught).

Next, an examination revealed a significant effect of report style, with more shallow MW[2] reported when self-catching than when probe-catching, $t(61) = 2.08$, $p = .04$, 95% CI [-.11, -.002], $BF_{10} = 1.04$, $d = .27$.

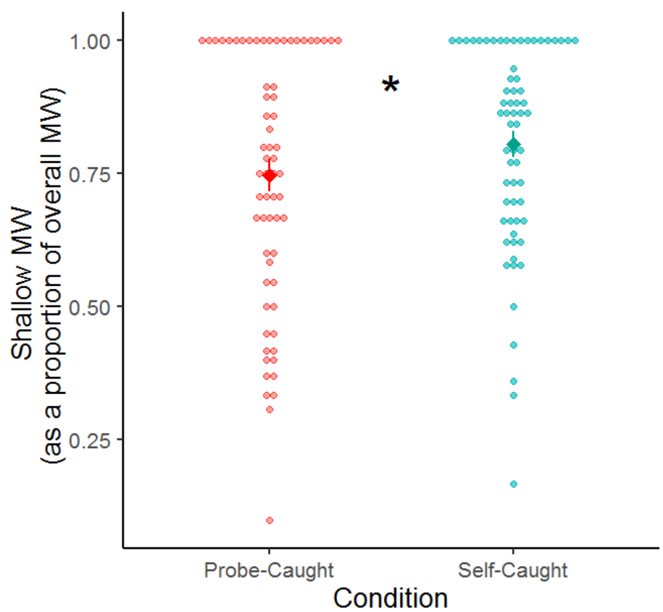


Figure 2.4. Shallow MW as a percentage of overall MW split by report type. Bars represent one standard error of the mean.

2.3.2.2 Retention performance, motivation and interest ratings

Retention performance did not differ based on report type $t(66) = .44$, $p = .66$, 95% CIs [-.08, .05], BF_{01}: 6.80, $d = .05$. Nor did report type influence motivation, $t(66) = 1.39$, $p = .17$, 95%

[2] Deep rates could have been examined, however as deep and shallow rates are ipsative, any results found for one would be identical – but in the opposite direction – for the other.

CIs [-.47, .09], BF_{01}: 2.99, d = .17 or interest ratings, $t(66) = .18$, $p = .86$, 95% CIs [-.30, .36], BF_{01}: 7.35, d = .03.

2.3.2.3 Perception of MW depth

Definitions of what constituted a deep MW instances typically contained reference to the instance being longer lasting, more elaborate thoughts, or leading to greater disengagement than shallow instances. Forty-six of 67 students (69%) reported that they interpreted the difference between shallow and deep MW instances in the same way between self-caught and probe-caught sessions.

2.3.3 Discussion

The present experiment provides yet further evidence that individuals can reliably become aware of and report self-caught MW instances. Furthermore, it demonstrates that MW is reported as being shallow significantly more often that deep during either self- or probe-catching. Given the high rate of shallow MW roughly (75% of instances), one possibility is that individuals are in a constant wavering, or never completely focused, attentional state. The 75-25% split in rates between shallow and deep reports are highly similar to those found for unintentional and intentional reports, and future work could benefit by examining whether unintentional MW is predominately perceived as being shallow (and vice-versa).

In contrast to my prediction, but supported by research on mindfulness, there were more shallow reports of MW in the self-caught condition than the probe-caught condition. This suggests that shallow MW is relatively easy to notice and that individuals have heightened meta-awareness to these subtle lapses in attention. However, that more shallow MW was noticed during self-caught reporting suggests that there may be underlying differences in report method and that the two report methods may tap into different mind wandering “types”. Given that mindfulness research suggests that individuals can be trained in mindfulness (Mrazek, Smallwood, & Schooler, 2012), an investigation of whether mindfulness training increases or decreases shallow self-caught MW could help to inform further benefits of mindfulness training. Although one of the benefits of using probes is that they are meant to provide a reliable and controlled way of measuring MW instances, it seems that the self-caught method is particularly effective at capturing shallow, fleeting MW instances. Note that this was contrary to what was predicted. And again, as in Experiment 2, MW report type did not influence retention performance; and a larger sample size would be needed to assess whether retention performance is predicted by MW depth. Overall, this experiment demonstrates a previously unrecognized benefit to the self-caught method, that despite prior suggestions, it is highly sensitive to MW, and furthermore, that MW can be categorized by participants as shallow or deep with high cross-participant reliability.

2.4 Experiment 4: MW characteristics

Experiments 2 and 3 examined whether differences exist in MW characteristics between self-caught and probe-caught reporting. In addition to intentionality and depth of thought, other factors have been reported to influence probe-caught MW rates (e.g., Task motivation: Antrobus, Singer, & Greenberg, 1966; Seli, Schacter, Risko, & Smilek, 2017; Unsworth & McMillan, 2013; Varao-Sousa & Kingstone, 2015; Task load: Forster & Lavie, 2009; Franklin, Mooneyham, Baird, & Schooler, 2014; Seli, Risko, & Smilek, 2016; Varao Sousa et al., 2013; Primed thoughts: Sayette et al., 2009). Through manipulations, researchers have influenced the frequency of MW reports, finding that participants tend to report less MW when motivated and less MW when task load is high. However, the influence of these factors has not been examined with the self-caught methodology, nor is it known whether other MW characteristics differ between the reporting types. As in the earlier experiments, determining whether self-caught MW differs fundamentally from probe-caught reporting in terms of the characteristics it captures is an important empirical question that could allow more diverse use of self-report methodologies aside from probe-catching alone.

To further explore potential differences between self-caught and probe-caught MW, the present experiment looked at the influence of a mental strain induction on MW rates and MW content. Specifically, I investigate how MW rates and content are influenced under manipulations of: 1) a mental strain induction (i.e., the requirement to give a task-related versus personal-related speech versus no speech control) and 2) the reporting method (probe-caught versus self-caught). One account of MW, the executive-control-failures account (Mcvay & Kane, 2010) suggests that since task-related executive-control processes are disengaged during MW, there is a consequence of deficits in primary-task performance, a hypothesis supported by numerous studies (Cheyne, Solman, Carriere, & Smilek, 2009; Christoff et al., 2009; Risko, Anderson, et al., 2012; Schooler et al., 2011; Smallwood, McSpadden, et al., 2007). The inclusion of a strain manipulation was included as a further "stress" to executive control processing. Numerous researchers have examined whether MW can be cued or primed by providing participants with intrusive thoughts or cuing personal goals (Baird et al., 2011; Banks & Boals, 2017; Masicampo & Baumeister, 2011; McVay & Kane, 2013; Stawarczyk, Majerus, Maj, Van der Linden, & D'Argembeau, 2011). In these experiments, MW tends to increase reportedly as a result of working memory and executive resource "overload", thus reducing the ability for participants to control MW occurrences (McVay & Kane, 2013).

In the present experiment, after an instance of MW is reported, participants were required to further categorize the event in terms of: task related/unrelated; positive/negative/neutral; how long the thought lasted; how freely moving the thought was; and emotional state. I predict that the task-related strain manipulation will increase required executive resources – as participants try harder to attend to the task – thus resulting in MW instances that are fundamentally different than the other two conditions. On the other hand, the personal-related manipulation will increase required executive resources, but these will be focused on the personal event under question and thus not draw increased resources to the task. More specifically, when individuals are under the task-related manipulation they will report increased task-related MW, whereas those in the personal-related manipulation will report increased task-unrelated MW, less freely moving thoughts and thoughts with greater emotional valence. As in the previous experiments, there is no theoretical evidence that self-caught and probe-caught MW instances differ, and thus I cannot hypothesize as to whether differences between the characteristics would occur in this series of manipulations. Speculatively, as suggested in Experiments 2 and 3, if certain types of MW are "easier" to notice these would be reported more frequently during the self-caught sessions than the probe-caught sessions.

2.4.1 Method

2.4.1.1 Participants

A 2 (report method) x 3 (strain manipulation) between-subject design was planned for the experiment with sample size was calculated using the following parameters in G*Power (Faul et al., 2007): alpha set to .05, power .8, effect size estimated at .25. This effect size was selected based on prior reports regarding the impact of writing manipulations on MW (Banks & Boals, 2017; Banks, Welhaf, Hood, Boals, & Tartar, 2016; Curci, Lanciano, Soleti, & Rimé, 2013).

These parameters indicated a required sample size of 211, with data collected for an additional month to ensure a sufficient sample was obtained after exclusions. Two hundred and seventy University of British Columbia undergraduate students participated in return for course credit, however 50 were removed from the analyses (41 due to no MW reported, thus eliminating the ability to examine differences in MW subcategories; and 9 due to technical program issues, i.e., video crashing). After exclusions, the sample included consisted of 166 women and 53 men (M_{age} = 20.8, SD_{age} = 4). Participants were pre-screened to include only individuals who reported being fluent in English, and normal or corrected-to-normal hearing and vision.

2.4.1.2 Materials

One 18-minute segment of a video-recorded lecture on population decline in Europe was used (same content as in Experiments 2 and 3).

2.4.1.3 Strain manipulation

Participants in the lecture-content speech condition were advised that at the end of the video lecture they would be asked to give a 5-minute summary speech on the content of the video. Those in the personal-content speech condition were told that at the end of the video lecture they would be asked to give a speech about a stressful occurrence that had occurred in the last six months and the impact it had on their life. Finally, those in the control condition were provided no further task instruction at this stage.

2.4.1.4 MW report method

MW report method was manipulated between-subject. For the probe-caught condition there were ten visual probes presented, with the amount of time between probes ranging from 60-185 seconds. Those in the self-caught condition were asked to indicate any off-task thoughts by pressing a labelled key on the keyboard. When a participant indicated that they off-task, they were then prompted to respond to the follow-up characteristic questions via keypress (Table 2.4), with only one response allowed per item.

Item question	**Item response options**
Was your thought:	Related the task; unrelated to the task
What type of thought were you having:	Positive; negative; neutral
How long did the off-task thought last:	0-5 seconds; 6-20 seconds; more than 20 seconds
On a five point scale, please rate how freely moving your thoughts were:	1 = focused on one topic – 5 = moving freely from topic to topic
Prior to reporting the off-task thought, how would you rate your emotional state:	Bored; frustrated; confused; engaged; flow state; happy; sad; not sure

Table 2.4. Summary of task means, with standard deviations in parentheses, with standard deviations in parentheses.

2.4.1.5 Retention test

Fifteen multiple-choice questions were created based on the video-lecture content, each with 4 choice options.

2.4.1.6 Mood ratings

Mood was assessed via a six-item PANAS subscale at the start of the experiment (before the manipulation was presented), and at the end of the experiment after the retention test. This assessment was conducted as a manipulation check to confirm that the strain manipulation had an effect. Participants responded on a 5-point Likert scale the extent to which they felt that emotion right then, where a response of “1” indicated very slightly/not at all and “5” indicated extremely.

2.4.1.7 Procedure

Participants provided informed consent before being brought into the experiment testing room. Prior to starting the experiment, participants were randomly assigned to one of the MW reporting conditions (self-caught or probe-caught) and one of the strain manipulation conditions (no manipulation, lecture-content speech, personal-stressor speech). Participants first completed the pre-experiment PANAS assessment before any task instructions were provided. Next, participants were advised that they would be asked to report their attentional state (probe-caught or self-caught instructions) while watching the video and that they would be completing a retention test at the end. Based on the condition, participants were provided no further instruction (control condition), told they would have to give an oral speech on the lecture content at the end of the task (task-related manipulation) or that they would be have to give an oral speech on a stressful event that had occurred recently (personal-related manipulation). At the end of the video, participants completed the retention test and a post-experiment PANAS assessment before being debriefed on the nature of the experiment.

2.4.2 Results

Table 2.5 displays descriptive information for each variable of interest, and the statistical outcomes are discussed in detail below. In cases where the samples were not homogeneous, as indicated by Levene's test, an adjusted degrees of freedom value is reported.

Measure	**Control condition**	**Task speech**	**Personal speech**
Mind wandering			
Probe-caught proportion (Range 0-1)	.24 (.21)	.27 (.17)	.28 (.16)
Self-caught count (Range 1-22)	6.12 (5.62)	4.15 (3.16)	5.83 (5.21)
Retention performance (Percentage)			
Probe-caught (Range 20-100)	64.38 (22.74)	68.48 (19.73)	72.22 (16.52)
Self-caught (Range 1-100)	68.43 (22.70)	76.88 (14.37)	66.67 (16.25)

Table 2.5. Summary of task means, with standard deviations in parentheses.

2.4.2.1 Overall mind wandering reports

Overall rates of MW were examined (as proportions for PC, and rates for SC) across speech manipulations. Probe-caught MW rates did not differ across the speech manipulation conditions, $F(2, 129) = .47$, $p = .63$, BF_{01}: 9.94, $\eta p^2 = .007$. Self-caught MW rates were also not impacted by the speech manipulations, $F(2, 97) = .84$, $p = .43$, BF_{01}: 5.08, $\eta p^2 = .017$.

2.4.2.2 PANAS ratings

Differences in PANAS ratings were examined across speech conditions to determine whether the manipulation had the desired effect of inducing a change in affect. If the manipulation had an impact, there should be a reported difference in affect rating from the start of the experiment (pre-speech information) as compared to the end of the experiment, based on the speech condition. Speech manipulation had no significant influence on positive affect scores comparing across start ($M = 8.05$, $SD = 2.08$) to end of the experiment ($M = 6.45$, $SD = 2.32$), ($F(1, 197) = .11$, $p = .90$ $\eta p^2 < .001$). Nor were negative affect scores impacted by speech manipulation when comparing start ($M = 4.77$, $SD = 2.09$) to end of the experiment reports ($M = 4.83$, $SD = 2.28$), regardless of speech manipulation, ($F(1, 197) = .58$, $p = .56$ $\eta p^2 < .006$). As the speech manipulation did not influence participant affect or overall MW reports, the following analyses do not consider this manipulation as a potential influencer on the MW responses.

2.4.2.3 MW categorization

For the categorization analyses, I examined the frequency in which each type of report was given and how this differed across MW report type by using a repeated measures design with report type as a between-subjects factor. Post hoc t-test analyses were adjusted with a Tukey correction for multiple comparisons.

Task relatedness. Task relatedness did not differ between self-caught and probe-caught reporting methods, $F(1, 211) = 3.25$, $p = .07$, $BF_{01} = .27$, $\eta p^2 = .015$, however for both report types participants were significantly more likely to report MW as task unrelated (64.2%) as compared to task related (35.8%), $t(211) = 6.79$, $p = < .001$, 95% CIs [-.36, -.20], $BF_{10} > 150$, $d = .45$.

Affect. With regard to affect, a between-subjects ANOVA indicated that the participant response did not differ between self-caught and probe-caught report condition: (positive affect) $F(1, 211) = .19$, $p = .66$, $\eta p^2 = .002$; (negative affect) $F(1, 211) = .009$, $p = .93$, $\eta p^2 < .001$; (neutral affect) $F(1, 211) = .038$, $p = .85$, $\eta p^2 < .001$. However, there was a significant effect of affect type, $F(2, 422) = 207.27$, $p < .01$, $BF_{10} > 150$, $\eta p^2 = .50$. Specifically, there was significantly more neutral MW (69.9%) as compared to either positive (10.7%), $t(212) = 18.99$, $p < .001$, 95% CIs [-.65, -.53], $BF_{10} > 150$, $d = 1.31$ or negative (19.3%) reports, $t(212) = 13.49$, $p < .001$, 95% CIs [-.58, .43], $BF_{10} > 150$, $d = .93$. Positive and negative reports also differed significantly from each other, with more negative than positive MW being reported, $t(422) = 2.74$, $p = .018$, 95% CIs [-.13, -.03], BF_{10}: 34.48, $d = .19$.

Thought length. Regardless of report type, MW instances were most likely to last under 5 seconds in duration (60.2%). There were significant differences in the amount of time spent MW, $F(2, 422) = 114.52$, $p < .001$, $BF_{10} = 6.78$, $\eta p^2 = .35$; this was further qualified by a time by report type interaction, $F(2, 422) = 7.59$, $p = .001$, $BF_{10} = 6.78$, $\eta p^2 = .04$. The comparison between self-caught and probe-caught reports, indicated that self-caught reports were more likely to occur in intervals under 5 seconds, $t(207.91) = 3.19$, $p = .002$, 95% CIs [.06, .25], BF_{10}: 11.91,

d = .44. On the flipside, probe-caught reports were more likely to last 6-20 seconds (t(209.32) = 2.25, p = .03, 95% CIs [-.18, .01], BF_{10}: 1.52, d = .31); or over 20 seconds (t(164.71) = 2.60, p = .01, 95% CIs [-.10, .02], BF_{10}: 2.87, d = .41) than were self-caught reports (See Figure 2.5).

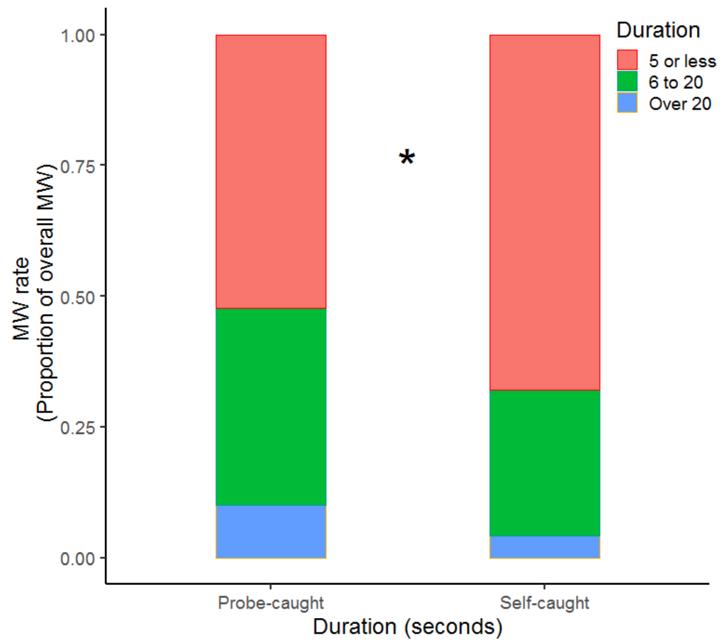


Figure 2.5. Mind wandering rates split by length, presented as a function of report type. Bars represent one standard error of the mean.

Freely moving thought. The degree of movement in MW reports was not influenced by reporting type, $F(1, 218) = .57$, $p = .44$, $BF_{01} = 12.5$, $\eta p^2 = .003$. However, there were significant differences in how freely moving MW was overall, $F(4, 872) = 17.76$, $p < .001$, $BF_{10} = 18.33$, $\eta p^2 = .075$. All conditions were significantly different from each other (all ps < .03), with the exception of thoughts that were completely focused on one topic and thoughts that were at the mid-point of freely moving. Participants were most likely to report thoughts as "Mostly focused" (M = .31) and least likely to report thoughts as freely moving (M = .09).

Emotionality. Emotion rating was not influenced by the method of reporting, $F(1, 211) = .84$, $p = .53$, $BF_{01} = 12.5$, $\eta p^2 = .004$. However, there were significant differences in the type of emotion reported, $F(6, 1266) = 103.45$, $p < .001$, $BF_{10} > 150$, $\eta p^2 = .33$, with participants most likely to indicate an emotional state of boredom (47.79%).

2.4.2.4 Retention test

Retention performance did not differ based on reporting method condition $F(1, 227) = .095$, $p = .75$, BF_{01}: 5.22, $\eta p^2 < .001$), speech manipulation $F(2, 227) = 2.45$, $p = .09$, BF_{01}: 3.98, $\eta p^2 = .02$, nor was there an interaction $F(2, 227) = 2.44$, $p = .09$, $\eta p^2 = .02$.

2.4.2.5 MW and retention test relationship

Probe-caught MW rates were significantly related to retention performance, $r(113) = -.29$, $p = .002$, with those reporting greater MW more likely to perform poorly on the retention test. On the other hand, self-caught MW rates did not relate to retention performance, $r(103) = -.18$, $p = .07$.

2.4.3 Discussion

In the present design I examined whether self-caught and probe-caught MW differ in the underlying characteristics of thought content (such as length and valence), and whether a speech manipulation would have an impact on reports. Unfortunately, a manipulation check indicated that the speech manipulation did not significantly impact participants' self-reported affect. In Banks and Boals (2017), where a future planning writing task was used to induce MW instances via primed personal goals, writing task condition did not influence MW rates. The authors suggest that this may have been due to limited processing time with regard to the writing task, and it is possible that in the present experiment, participants did not have enough time to process the speech manipulation. Furthermore, the absence of a difference in PANAS scores across conditions suggests that the present manipulation may not have been powerful enough to induced task strain as expected. Despite the lack of influence of manipulation, results from the present experiment indicate that some aspects of wandering thoughts may be contingent on the way that MW is reported, and provide further information about the content of MW reports.

While completing an in-lab video lecture task, the method of MW response (probe-caught versus self-caught) rarely had a differential impact on all five of the factors considered: task relatedness, affect, length, free movement and emotionality of the wandering thought. Specifically, only thought duration differed across report type with self-caught reports significantly more likely to be shorter in duration than probe-caught reports. Thus it may be the case that self-caught thoughts are fleeting occurrences that one is quickly able to capture and report. That said, participants did report that probe-caught instances were on average shorter than longer, thus suggesting that the probe was catching them just in the earlier stages of mind wandering. Both self-caught and probe-caught reports were more likely to be task-unrelated than task-related. This was an unexpected outcome, as it might be thought that a task-unrelated thought may be more likely to trigger a self-caught report – since it would be so divergent from the task, thus acting as a cue to the participant to report the instance. Both probe-caught and self-caught MW rates were most likely to be reported as neutral in content, contrary to some research suggesting that MW relates to ruminative processes (Levinson, Smallwood, & Davidson, 2012; Ottaviani et al., 2015; Ottaviani, Shapiro, & Couyoumdjian, 2013; Smallwood, O'Connor, Sudbery, & Obonsawin, 2007). When examined how freely moving participants' thoughts were, in both instances, MW rates were more likely to be focused on one topic than freely moving. The emotional tone of the wandering thought was most likely to be boredom for roughly 50% of both types of reports. The finding that probe rates were similar to self-caught rates in many of the qualitative cases suggests that using either of the two self-report methods can provide valuable feedback related to MW content. The impact of MW reports on retention test performance mirrored past findings that retention performance is worse for higher rates of probe-caught MW, and that self-caught MW does not appear to be related to retention performance.

This experiment provides further support that the self-caught MW methodology is a reliable way to acquire MW data. Despite most MW research being conducted using the probe-caught measurement, this experiment suggests that individuals are highly capable of noticing their minds wandering (with up to 22 reports in a 20 minute task), and furthermore, self-caught instances of MW are often similar to those being reported via the probe-caught method. The next experiment examines self-caught and probe-caught MW within a more natural setting, a classroom lecture, and examines whether the addition of probes impacts the rate of self-caught reports. This direct comparison of methods will determine whether self-caught MW rates are altered by the introduction of thought probes, with such a divergence indicating that the probe-caught method is creating an artificial testing environment.

2.5 Experiment 5: MW reporting in classrooms

MW has been reported to occur regularly within classroom settings, as assessed by probe-caught reporting (Lindquist & McLean, 2011; Risko, Anderson, Sarwal, Engelhardt, & Kingstone, 2012; Varao-Sousa & Kingstone, 2015; Wammes, Boucher, Seli, Cheyne, & Smilek, 2016; Young, Robinson, & Alberts, 2009). However, as the previous experiments in my thesis have demonstrated, individuals are also quite capable of noticing lapses of attention without utilizing external probes. The results also suggest that the MW events are often qualitatively similar between self-caught and probe-caught instances, although there are exceptions – perhaps most notably in terms of the fact that probe-caught MW relates to retention deficits whereas self-caught MW does not (as per Experiment 4). The experiments examined thus far have measured MW using naturalistic learning tasks (reading and lecture viewing) within a laboratory setting. To build on my earlier experiments, I examined whether rates of self-caught and probe-caught MW and their impact on retention converge, or diverge, when examining their occurrence in more complex real-world naturalistic settings.

The aim of the present experiment was to examine self-caught MW rates in a real learning setting, and also test to determine if self-caught MW rates change when the probe method is introduced. Specifically, it is plausible that the experimental method of probing individuals and asking them to reflect on their state of mind creates an artificial situation that is influencing MW results. By using the self-caught results as the dependent variable, and manipulating the presence of probes, I can determine whether self-caught MW reports are influenced by the introduction of probes. Given the exclusive use of the probe-caught method to measure MW within classroom settings, the implied assumption is that the probe-caught method provides an objective measure of MW and, if true, self-caught MW rates would be unaffected by this “artificial” method. However, given the greater ecological validity of the self-caught method, examining rates within a natural setting might reveal divergent outcomes. Even if the probe-method is a reliable and non-influencing means for collecting MW reports, I hope to further demonstrate the ease at which the more natural self-caught method can be used in every day settings.

Though the consistency of probe-caught MW results speak to both the stability of MW, and the conclusion that the lab results scale up to real life situations, one might ask whether this stability is due in part to the use of the thought-probe methodology itself. In short, is interrupting individuals and asking them to reflect on their state of mind creating a common, artificial situation at all levels of investigation and *this* is the reason for the reliable MW results? Or to put it even more forcefully, the robust MW results may be anything but evidence that MW is relatively stable across different settings; rather it is evidence that the probe-based methodology is inserting a singular common artificial event into different situations. Accordingly, the stability of MW may merely be an illusion that reflects the common practice of using the probe-based method to measure MW. If true, then probes not only influence the specific moment of query within a task, but how participants view, and in turn perform, during the entire experiment. The aim of the present study is to address this issue by examining MW using the ecologically valid self-caught methodology, and then testing if self-caught MW rates change when the probe-based methodology is inserted into the experiment.

The present experiment examined students' ability to self-catch MW during three classroom sessions. Thought-probes were introduced in the middle session only. By comparing MW rates in the self-caught sessions (sessions 1 & 3) with the MW rates in the self-caught/probe-caught session (session 2), I could examine if self-caught MW rates are altered by the introduction of thought probes. If the hypothesis is correct, and the probe-caught method is creating an artificial testing environment, then self-caught MW rates in sessions 1 and 3 will be significantly different from session 2 when MW probes are included. Alternatively, if the probe-caught method is providing a 'pure measure' of MW (as the field assumes), then the self-caught MW rate should be relatively stable across all three sessions and, most crucially, be unaffected by the introduction of thought probes in session 2.

2.5.1 Method

2.5.1.1 Participants

Participants in this experiment were students enrolled in an Introductory Psychology course at the University of British Columbia (UBC). The students (N = 259) were informed of the testing dates and general task protocol via the UBC Learning Management System (Connect), however only data for participants who completed all three sessions of the experiment are included in the analyses. Three additional participants were removed from analyses for reporting self-caught MW rates greater than 3 SDs above the mean (thereby influencing kurtosis and skew of the data). Of the included participants (n = 86), sixty-three were female, and participant age ranged from 17-28 years (M = 19.07, SD = 1.72). Participants provided informed consent before each session and received course credit for each session they completed. In the present experiment power analyses were not conducted as sample was restricted to those registered in the class, and those who attended the classroom session on the testing days.

2.5.1.2 Course details

The course was offered from 9:30-10:50AM on Tuesdays and Thursdays. The testing sessions covered content related to Neurons and Neurotransmitters (session 1); Genetics and Twin Studies (session 2); and Heredity and Evolution (session 3). Participants were provided a response sheet at the start of each class to record their answers.

2.5.1.3 Self-caught mind wandering

In all three sessions participants were asked to self-catch MW and report any instance by placing a check mark in the 10-minute interval that corresponded with the current time (e.g., 10:10-10:29) on their response sheet. Laboratory experiments investigating self-caught reports tend to have participants respond via keypress whenever MW is noticed. As students were not provided with electronic devices as part of this class the 'paper and pencil' method was a natural solution.

2.5.1.4 Probe-caught mind wandering

In session 2 only, participants responded to six visual probes in addition to reporting self-caught MW. Probe frequency was selected based on prior research in live lecture environment (Lindquist & McLean, 2011; Varao-Sousa & Kingstone, 2015; Wammes et al., 2016). Probes were presented on PowerPoint slides that were incorporated into the class lecture. The probes asked students to indicate whether they were on task or MW by circling their response on their response sheet. Participants were given roughly five seconds to answer each probe.

2.5.1.5 Retention test

At the end of each lecture six multiple choice test questions were displayed to the class via PowerPoint, with each question having four choice options. Test questions were created after previewing the lecture slides from the course instructor. Participants indicated their answer by circling the option they felt to be correct on the response sheet. Participants were provided 10 minutes to record their answers and were asked to not use their class notes or consult with classmates during the testing period.

2.5.1.6 Motivation and interest ratings

Once completed, participants reported their motivation and interest in the lecture ("How motivated were you to attend to the lecture?" and "How interesting did you find the material presented in today's lecture?"). A 5-point Likert Scale was used, where 1 = Low Motivation/Interest and 5 = High Motivation/Interest. These responses were indicated by circling the number that participants felt corresponded to their experience. These measures were collected to replicate past findings that both factors relate negatively to probe-caught MW rates

(R. Ben Hollis & Was, 2009; Lindquist & McLean, 2011; Unsworth & McMillan, 2013), and to examine if this relationship extends to self-caught MW.

2.5.1.7 Prior experience

Lastly, participants recorded how much experience they had with the content presented in the lecture. Introductory psychology classes at UBC tend to draw in many upper year students from other areas, with these students having prior exposure to similar content (e.g., Neuroscience or Biology majors). Participants were asked to indicate experience with the material by circling one of three options: I’ve taken multiple courses on the topic; I’ve taken 1-2 courses on the topic; This was my first lecture on the topic. I suspected that this factor may influence retention test performance.

2.5.1.8 Procedure

At the start of each class period a link to an online consent form was made available to students. Any student wishing to participate was asked to go to the e-form to provide consent. Students were provided response sheets on which all responses were made. On the response sheet, students recorded a numeric identifier that allowed researchers to anonymously track performance over the 3 sessions. The following instructions were provided verbally at the start of each session: “During the lecture I’d like you to note any mind wandering that you experience by putting a checkmark in the box that corresponds to the current time. This means that you might have some time block (for example: 10:00-10:09) where you have 8 check marks, some where you have 2 or 3 and some where you have none. Simply put a checkmark each time you notice mind wandering, in the corresponding time block. Mind wandering is any thought that is not related to the course material being presented. Examples of mind wandering include thinking about what you are going to have for lunch, thinking about something you did on the weekend, other course work, etc." At Session 2, participants were also provided the following instructions: "Additionally, you will be asked to indicate your attentional focus at a number of specific time points during the lecture. During the lecture you will see a power point slide that asks you to indicate whether you were mind wandering or on task. When you see this slide, again on the sheet provided, simply circle your response to indicate your thoughts in the moments before you saw that slide. Remember you provide reports both anytime you notice mind wandering and at the specific time points where the question is on the slide". Students were reminded that participation was optional, and that their instructor would not access individual responses or know whether students chose to participate. Participants were then given the opportunity to ask questions, after which the lecture began. Lectures ran roughly 60 minutes, with 10 minutes provided at the end of each session for participants to complete the retention test and follow-up questions.

2.5.2 Results

Across all three sessions a small majority of participants (54%) reported having taken 1-2 courses on the topic presented[3]. Descriptive statistics are presented in Table 2.6.

Measure	Session 1	Session 2	Session 3
Self-caught frequency (Range 0-23)	6.72 (5.10)	6.33 (4.43)	5.97 (3.35)
Probe-caught proportion (Range 0-1)	-	.40 (.23)	-
Retention test score (Range 0-1)	.73 (.18)	.91 (.17)	.81 (.18)
Motivation rating (Range 1-5)	3.66 (1.00)	3.11 (.96)	3.13 (.95)

[3] A between-subjects analysis with 'Experience' as a factor indicated that none of MW rates, Interest or Motivation ratings were impacted by differences in prior experience with the lecture topic (all *p*s > .12). Memory test performance, however, was impacted such that students with prior experience performed significantly better compared to those with no prior experience, *p*s < .04.

Interest rating (Range 1-5)	3.32 (.79)	3.01 (.79)	2.81 (.73)

Table 2.6. Task means, with standard deviations in parentheses.

2.5.2.1 Self-caught mind wandering reports

A repeated measures ANOVA suggested that self-caught rates of MW did not differ significantly across the 3 sessions, $F(2, 170) = 1.44$, $p = .24$, $BF_{01} = 13.56$, $\eta p^2 = .02$. This BF indicates that the data are 13.56 times more likely under the null hypothesis (that there is no difference across sessions) than under the alternative hypothesis[4].

2.5.2.2 Probe-caught mind wandering reports

In Session 2 participants indicated that they were MW in response to thought probes 40% of the time. This rate falls squarely in the centre of the predicted 30-50% range based on a vast wealth of past work (Lindquist & McLean, 2011; Risko, Buchanan, Medimorec, & Kingstone, 2013; Unsworth, McMillan, et al., 2012; Varao-Sousa & Kingstone, 2015; Wammes, Boucher, et al., 2016)

2.5.2.3 Mind wandering time course

As can be seen in Figure 2.6, self-caught MW reports were blocked by the 10-minute interval in which participants responded (e.g., 9:40-9:49). Analyses indicated that there was a significant difference in self-caught reports across session 1: $F(4, 340) = 3.42$, $p = .009$, $BF_{10} = .10$, $\eta p^2 = .04$; but not for session 2: $F(4, 340) = 1.48$, $p = .21$, $BF_{01} = 56.96$, $\eta p^2 = .02$; or session 3: $F(4,$

[4] It is conceivable that completing only session 2 would lead to performance reports that differed from those who also had the chance to 'ease into the study' with session 1. That is to say, perhaps probe independence was a product of participants completing a session without probes (i.e., session 1) which in turn influenced session 2 performance. One way to test this possibility is to compare self-caught MW rates at session 2 for those who only completed session 2 with those who had also completed session 1. These data were limited to only 12 participants and so a full analysis could not be completed due to small sample; however session 2 participants who did not also complete session 1 reported self-caught MW at a rate ($M = 6.5$, $SD = 4.56$) that was very similar to those who did both sessions ($M = 6.33$, $SD = 4.43$).

340) = 0.78, p = .54, BF_{01} = 84.42, ηp^2 = .009. For Session 1, a linear trend analysis was conducted to determine if self-caught reports increased over time, however this was non-significant, $F(1, 85) = .37$, $p = .55$, $\eta p^2 = .004$.

Figure 2.7 presents the time course for probe-caught MW responses. A repeated-measures ANOVA revealed a significant effect of time on task, $F(5, 425) = 3.97$, $p = .002$, BF_{10} = 3.81, ηp^2 = .05. A linear trend analysis revealed a significant linear trend, $F(1, 85) = 5.77$, p = .018, ηp^2 = .06, in line with prior work suggesting that MW increases over time on task (Farley et al., 2013; Lindquist & McLean, 2011; Seli, Wammes, et al., 2015). Follow up analyses indicated that, after Bonferroni corrections, only Time 2 was significantly different, such that less MW occurred there, than at Time 3 or Time 6, $ps < .04$.

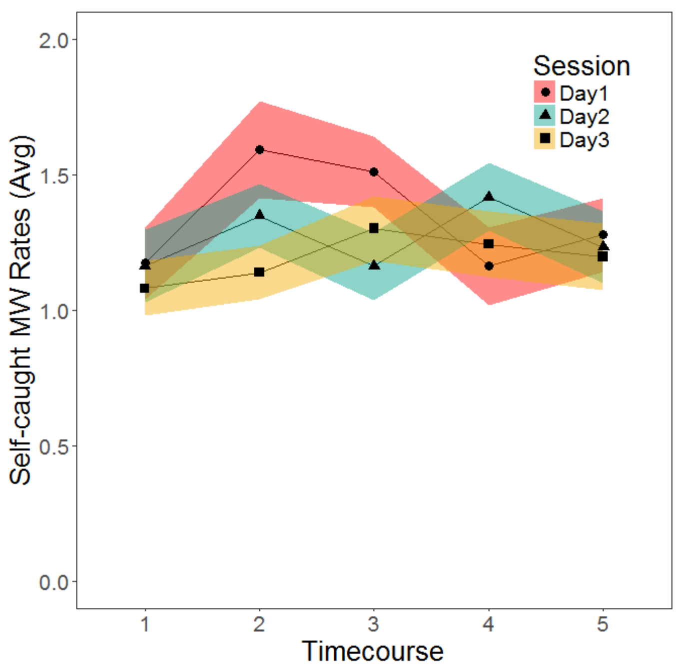


Figure 2.6. Time course of self-caught mind wandering reports split by session. Coloured bands represent one standard error of the mean.

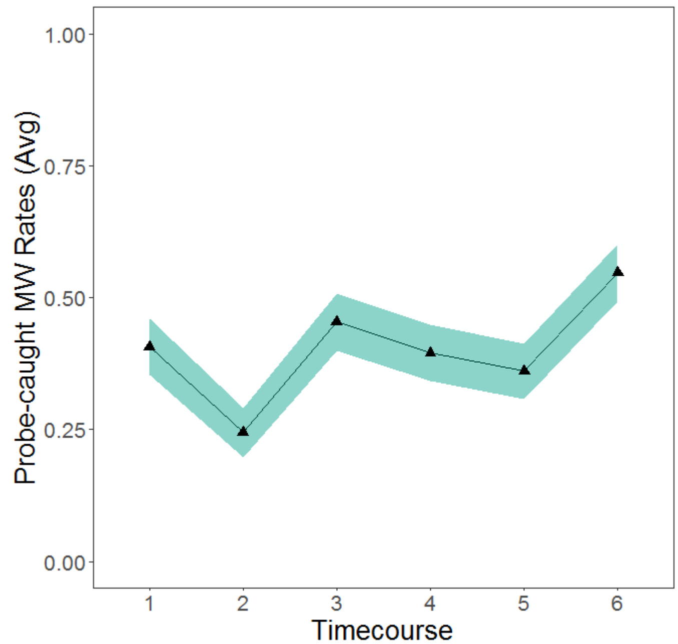


Figure 2.7. Time course of probe-caught mind wandering reports in session 2. Coloured bands represent one standard error of the mean.

2.5.2.4 Retention test

A repeated measures ANOVA revealed that retention test performance differed significantly across the three testing sessions, $F(2, 170) = 33.80$, $p < .001$, $BF_{10} > 150$, $\eta p^2 = .29$. Post-hoc analyses using a Bonferroni correction for multiple comparisons (corrected p-value: .05/3 = .017) revealed that retention test performance differed significantly between each of the conditions (all *ps* < .002) with the best performance in Session 2 and the worst performance in Session 1.

2.5.2.5 Correlations

Mind wandering rates

Table 2.7 displays the statistical relationships between MW reports across all sessions. Individuals' MW rates were significantly and positively correlated across all sessions, with the exception of session 1 self-caught reports with session 2 probe-caught reports ($p = .07$), although the relationship was still positive. This suggests that the ability to report MW remained consistent at a participant level across sessions.

Sessions	Session 2 (self-caught)	Session 3 (self-caught)	Session 2 (probe-caught)
1 (self-caught)	.63***	.53***	.20
2 (self-caught)	-	.53***	.50***
3 (self-caught)	-	-	.27*
**p < .05. **p < .01. ***p < .001*			

Table 2.7. Pearson's r correlations for mind wandering measures across all three testing sessions. (*N* = 86).

Mind wandering and other measures

Table 2.8 displays statistical summaries for the correlations between MW with the other measures collected. Across all three sessions no relationship between MW and retention test performance was found. This was the case for both self-caught MW reports, and probe caught-MW reports. Both self-caught and probe-caught MW reports were significantly negatively correlated with Interest and Motivation ratings. This relationship suggests that as interest and motivation in the course content decreases, MW reports increase.

	SESSION 1			SESSION 2			SESSION 3		
MW	Retention	Interest Rating	Motivation Rating	Retention	Interest Rating	Motivation Rating	Retention	Interest Rating	Motivation Rating
Self-caught	.08	-.24*	-.31**	.13	-.25*	-.31**	.20	-.29**	-.36**
Probe-caught	-	-	-	-.06	-.23*	-.52**	-	-	-
**p < .05. **p < .01*									

Table 2.8. Pearson's r correlations for across all three testing sessions.

2.5.3 Discussion

In a natural classroom setting, the addition of probe-caught reporting did not influence the time course or rate of self-caught MW, although it may have influenced how students performed on a later retention test. The finding that probe rates were similar to those found in prior classroom research, and did not affect self-caught MW rates, strongly suggests that the past consistency of probe-caught MW rates across a range of different settings is not an artifact of the thought-probe method. This experiment also indicates that the self-caught MW methodology is a reliable way to acquire MW data. The extension of measurement techniques to include students' self-caught reports provides valuable information on how to successfully and naturalistically monitor MW in lecture settings, outside of the laboratory.

This experiment investigated if the relatively stable MW rates found across past investigations is an artifact of the prevalent use of the thought-probe method. The obtained data reject this hypothesis. In the present experiment, the probe-caught MW rate was 40%, which dovetails with the rates reported during other lecture-based experiments (Lindquist & McLean, 2011; Varao-Sousa & Kingstone, 2015; Wammes et al., 2016). If the inclusion of thought-probes was altering the testing situation, the self-caught MW rates in session 2, where thought-probes were also present, should have differed from sessions 1 and 3 where only self-caught reports were collected. However, neither the self-caught MW rates, nor their time course across the three lectures differed. Students caught themselves MW roughly 6-7 times per lecture, with these reports evenly distributed over time. Thought probes in session 2 had no impact on self-caught performance. Thus these data indicate that the probe-caught method is a valid sampling method, and the stability in MW that this method yields is valid and not an artifact of experimental design.

In addition to recording MW, I measured retention test performance, interest and motivation ratings. As each session covered a different topic it is not surprising that I found variation in terms of these factors. Perhaps more interesting is that although these subjective ratings varied across the different sessions, self-caught MW rates did not. This suggests that individual differences in attention are more stable than subjective ratings of interest or motivation. MW rates were also highly and positively correlated, speaking further to the stability of individual MW variation.

I also replicated prior work reporting a relationship between probe-caught MW with both Interest and Motivation ratings, and found evidence that this finding extends to self-caught MW rates (Forster & Lavie, 2014; Lindquist & McLean, 2011; Unsworth & McMillan, 2013). And I replicated the recent finding that MW may not correlate with retention test performance in live lectures (Varao-Sousa & Kingstone, 2015; Wammes, Seli, et al., 2016: who found that type of MW moderated the impact on short term versus long term retention), though it is worth noting that as overall retention performance was quite high it is possible that ceiling effects reduced the ability to detect a correlation. Also noteworthy is that retention performance was highest in Session 2, where both self-caught and probe-caught monitoring were occurring. It is conceivable that the presence of both methods raised meta-awareness and task ability, thus resulting in better performance; however that MW rates were not impacted by this combined reporting experience suggests increased *attention* is not responsible for the Session 2 retention increase. An alternative explanation is that the material presented in Session 2 was simply easier to remember.

Prior MW research has been dominated by probe-caught measurement. Although researchers have proposed that individuals lack enough meta-awareness to reliably report MW without a thought-probe (Jackson & Balota, 2012; Sayette et al., 2009; Schooler et al., 2004), the findings of this experiment suggest that individuals are quite capable of noticing and reporting MW during a natural lecture setting. The use of the self-caught method provides stable rates of MW in a live lecture setting, indicating that this is a viable method for future research.

Furthermore, as demonstrated in the present experiment, self-caught MW rates are unaffected by the introduction of the probe-caught method, meaning that the ecologically valid self-caught method can be used alongside the more controlled probe-caught method without any negative effects and with the added benefit of collecting additional data. In short, these methods appear complementary and by availing themselves to both, MW researchers can have the best of both worlds.

2.6 Chapter summary

Together, the experiments presented in Chapter 2 examined whether self-caught and probe-caught methods capture similar or distinct mind wandering experiences. The results from experiments 2-5 suggest that both reporting methods (probe or self-caught) are efficient ways of capturing a MW instance. In many cases, the qualities or subtypes of the instances are similar and occur at similar frequencies (i.e., intentionality, affect, thought movement) however in other cases MW reports differ based on their quality (i.e., depth, duration). That is to say, depending on the type of wandering thought one is hoping to capture, it appears that the two self-report methods could be used interchangeably, and even combined.

The first experiment in this chapter provided evidence that the way in which a MW prompt is presented (i.e., audio or visual) does not impact MW reports. This check allowed me to move forward to using a MW probe method that was best suited to a given task, and not be constrained to one that matched the primary task modality. In Experiment 2, I investigated how probe-caught and self-caught MW rates were impacted during two naturalistic learning tasks (lecture watching and non-fiction text reading), and whether MW intentionality was influenced by task or report method. In line with past probe-caught research, and extending to self-caught methods, unintentional MW was reported more frequently than intentional MW – a finding that was stable across both learning tasks. Interestingly, there was significantly more unintentional MW reported when reading than when watching. This outcome suggests that individuals are more likely to *choose* to disengage (i.e., intentionally MW) while watching than while reading – even though material watched was rated as more interesting. The results are in line with Seli et al. (2016) who found higher rates of MW intentionally during easy versus difficult tasks, if one considers watching a video lecture to be "easier" (or less cognitively demanding) than reading.

Results from Experiment 3 suggest that MW is more often shallow than deep in nature, regardless of the instance is self- or probe-caught. There was also evidence for more reports of shallow MW when self-catching than when probe-catching, conversely suggesting that probes are more likely to capture deep instances of MW. The greater frequency of shallow reports when self-catching may reflect the fact that individuals are often on the cusp of MW or have fleeting thoughts regularly "pop" into their minds, and the freedom to report at any instance captures these occurrences. The abrupt probe onset may trigger individuals to report an instance as deep more often – as the thought could not be so fleeting or shallow if they were caught by an external probe before realizing it.

Experiment 4 was designed to further build on the findings of Experiments 2 and 3 and determine whether self-caught and probe-caught reports were capturing different experiences of MW in terms of thought valence, relatedness, duration, movement or emotionality. Unfortunately, the goal of manipulating thought content via a speech manipulation condition was unsuccessful and not a powerful enough motivator to change participants' affect. Nevertheless, the experiment was still able to provide information on categorizing MW thoughts and determined that many of the characteristics of MW do not differ between the two methods.

Specifically, it seems that under both reporting styles MW reports are most likely to be unrelated to the task, neutral in content, last less than 5 seconds, and occur alongside feelings of boredom. In line with prior work, increased probe-caught MW rates predicted poorer task performance, whereas self-caught reports were not predictive of performance.

The goal of Experiment 5 was to test whether the probe method creates an artificial environment, thus impacting inattention reports, as well as to extend the use of self-caught and probe-caught methods to the classroom as this is a non-artificial, natural environment where MW frequently occurs (Lindquist & McLean, 2011; Risko et al., 2012; Unsworth, McMillan, Brewer, & Spillers, 2012; Varao-Sousa & Kingstone, 2015; Wammes et al., 2016; Young et al., 2009). There was no influence of probe-caught reporting on the time course or rate of self-caught MW, although the inclusion of both probes and self-caught reporting may have positively influenced how students performed on a later retention test.

Collectively the data from this chapter converge on the conclusion that both self-caught and probe-caught methods are valid techniques for measuring MW. Does this mean that each method could be used interchangeably or combined seamlessly without one method influencing the other? At one extreme, the data of Experiments 2-5 could indicate that each method is capturing a similar type of attentional lapse. This follows if the self-caught method provides participants the opportunity to indicate MW that occurs between probes, notwithstanding the fact that instances may still be "caught" at times when participants have not yet realized that they are mind wandering (Jackson & Balota, 2012; Sayette et al., 2009; Schooler et al., 2004). It is also supported by the overlap in categories or features of MW described in Experiments 2-4, as self-caught and probe-caught rates were similar across the majority of the MW characteristics. At the other extreme it is conceivable that these two MW methodologies do not interact because they are capturing distinct types of MW (e.g., probes are measuring lapses in attention that operate outside of conscious awareness). One point of support for self-caught and probe-caught MW being distinct measures of MW can be found in the dissociated relationships reported in Experiment 4. Specifically, probe MW rates predicted retention performance rates while self-caught did not, and the duration of a thought was more variable under probe-caught than self-caught reporting (although this dissociation was not found in Experiment 5). Further tentative support for this idea is found in the different time course data for the two types of MW methods as examined in Experiment 5, such that probe-caught MW showed a linear trend, increasing over the course of the lecture, while no linear trend was found for self-caught MW. It is worth highlighting that although probe-caught MW showed a linear trend, this was driven by only 2 time points, and thus the stability of the pattern is unclear. The lack of strong support is consistent with research suggesting that in live lectures attention may wax and wane to a different rhythm than in laboratory settings, and thus warrants further exploration (Wammes, Boucher, et al., 2016; Wammes & Smilek, 2017).

One strong argument for use of the probe-caught method is that probes will *always* catch a mind wandering instance, regardless of its nature; however, giving individuals the freedom to self-report MW appears to capture a similar distribution of thought. Across experiments 2-5 the relatively high rate of self-catches (Mean range across experiments = 5.5 - 15.76) suggests that researchers may be missing out on regularly noticed MW instances when restricting use to the probe-caught method. Given that report style did not appear to influence MW rates at the overall level, the following experiments used a probe-caught methodology as it was somewhat easier to implement and allowed for greater generalizability to the existing literature. In the next Chapter of this thesis, I continue an examination of MW within naturalistic settings, focusing on classroom lectures and the leisurely activity of playing video games.

Chapter 3: Measuring MW in Naturalistic Settings

The previous chapter provided strong evidence that both MW methodologies (probe-caught versus self-caught) provide reliable measurements of MW performance. In the present chapter (Experiments 6-8), I use the probe-caught reporting method to examine MW rates in a variety of naturalistic settings. Experiments 6 and 7 manipulate features of the classroom setting (i.e., instructor visibility and peer presence) to investigate the impact on student learning and MW rates. Experiment 6 investigates whether retention and MW differ between a traditional classroom and a classroom where the lecture is presented via pre-recorded video. Specifically, how does the presence of a video recorded professor influence MW rates and retention, relative to a live professor? Experiment 7 manipulates lecture video display format and social learning settings to understand their influence on student learning and MW experiences. Thus, this experiment asks: what are the main effects of presentation and setting on MW and on comprehension? And, does social setting or instructor visibility moderate the relationship between MW and content retention? In Experiment 8, I examine MW in the context of a naturalistic and enjoyable leisure setting, playing a video game. This experiment examines the question of what the potential “lower bound” on MW in a natural task might be.

3.1 Experiment 6: The impact of lecture format on MW and learning

Classroom and learning settings are reported to be the most common environments for fostering MW (Kane, Gross, Chun, Smeekens, Meier, Silvia, & Kwapil, 2017; Unsworth & McMillan, 2017; Unsworth, McMillan, et al., 2012). In the past decade, online learning environments have become a staple in both university settings and in open forums (e.g., edX, Coursera, etc.). At Canadian universities roughly 20% of students enroll in online courses

(Canadian Digital Research Learning Association, 2019; Canadian Virtual University, 2012). As hundreds of thousands of students from around the world sign up for online courses there has arisen the opportunity for large quantities of data to be collected—these data have focused on lecture length, drop-out rates, accessibility, participation, overall sense of satisfaction and interest (Grainger, 2013; Perna et al., 2014; Wuensch, Aziz, Ozan, Kishore, & Tabrizi, 2008). Yet little research has been conducted on whether real-time attention and learning are allocated differently between live and video learning environments. While the opportunity for accessible and continuous learning is unquestionably important, key questions have been raised: how do different learning environments impact the learning experience? Specifically, are there cognitive costs or benefits to changing the live lecture experience? And what is the subjective student experience of different learning environments? Answering these questions requires a detailed investigation of what happens when the traditional, live classroom-based learning environment changes to one that is video-based. The present experiment explores potential differences in attention and retention that may arise when learning shifts from a live learning environment to one that is entirely video-based.

The research that has been conducted on in-class learning as compared to video or e-learning has resulted in contradictory outcomes (Abdous & Yoshimura, 2010; Demetriadis & Pombortsis, 2007; Ferguson & Tryjankowski, 2009; McKinney, Dyck, & Luber, 2009; Sankaran & Bui, 2001). The discrepancy in results could be confounded by the many differences between settings, since direct comparisons between live and video lectures does not allow for a systematic separation of factors that might differentially impact learning (e.g., learning setting, peer presence, professor presence, content visible, etc.). Without controlling or equating these other factors the reason for a change in performance, or lack-there-of, cannot be isolated. This lack of thorough investigation is noteworthy given that many universities offer blended-learning courses that often have a heavy video-viewing component, with the intention that these changes are providing a beneficial change to the educational experience. Furthermore, many of these experiments have taken place in lab environments which do not adequately reflect student motivation or performance compared to real settings.

The present experiment investigated the impact of lecture presentation style on content retention and MW reports. Specifically, I focused on *one* factor that differs between live and video-learning settings: a professor's physical presence. The reason for investigating this factor was that it could be manipulated while holding constant all other factors in the session (e.g., time and location of lecture, expectations, peer presence). In recent years, as video and online learning have become more prominent, researchers have examined how online engagement is influenced by instructor presence (Kovanović et al., 2018; Yildirim & Kilis, 2019; Zhang, Lin, Zhan, & Ren, 2016). By manipulating only the "live" quality of the professor and maintaining all other aspects of the classroom setting (e.g., scheduled time and location, presence of peers, etc.) I can begin to isolate what aspects of the lecture, live versus video format, impact the learning experience. In the following pages I review how these relate to past investigations of retention and MW in classroom settings.

To gain an understanding of how different learning environments impact the learning experience, and whether there are cognitive costs or benefits when the classroom setting is changed, I must first consider which cognitive factors are likely to be impacted. To date, researchers have focused largely on retention performance when comparing classroom environments. However, studies have yielded mixed outcomes ranging from: equivalent performance in technology-based learning and live classrooms (Abdous & Yoshimura, 2010;

Demetriadis & Pombortsis, 2007; Sankaran & Bui, 2001), to benefits for e-learning (McKinney et al., 2009), to benefits for classroom learning (Ferguson & Tryjankowski, 2009; Xu & Jaggars, 2013). Methodological and/or situational differences within and between experiments have varied dramatically, making it unclear what, if, and how specific factors are responsible for the different retention effects. In contrast, my experiment strives to hold all variables constant, save one—the nature of the instructor's presence.

Another factor that may be impacted by changes to the classroom environment is MW. Indeed, a daily diary study found educational settings to be the most common location for MW to occur (Unsworth, McMillan, et al., 2012), which is convergent with reports that MW in classrooms occurs 30–50% of the time (Cameron & Giuntoli, 1972; Lindquist & McLean, 2011; Schoen, 1970; Varao-Sousa & Kingstone, 2018; Wammes, Boucher, et al., 2016). While these experiments provide information on how attention may lapse during a live lecture setting, no experiments to date have investigated whether the amount of MW changes as a function of classroom setting (but see: Wammes & Kingstone, 2018 for changes in degree of MW across learning settings). What limited research has been conducted in classroom styled environments suggests there is a negative correlation between MW reports and retention performance (Lindquist & McLean, 2011; Risko, Anderson, et al., 2012; Szpunar, Khan, & Schacter, 2013; but see also Varao-Sousa & Kingstone, 2018; Wammes et al., 2016). Given that MW can negatively impact retention performance I reasoned that both factors should be investigated when considering a change in an alternative learning environment.

To date, fundamental methodological differences between learning environments prevent a straightforward comparison between the performance effects of live and video-based learning situations. For instance, students in e-lecture environments can re-watch material and are often encouraged to do so (Demetriadis & Pombortsis, 2007; McKinney et al., 2009). Additionally, the environment where learning takes place is often vastly different (e.g., professor presence, material modality, social setting, location of viewing, etc.) making it difficult to pinpoint what factor, or combination of factors, translates into differences in performance. By systematically manipulating individual factors of the learning environment one can begin to isolate what cognitive changes, if any, emerge in different settings. The present investigation manipulated only one factor, physical professor presence, while holding all else constant, allowing for a direct comparison between learning environments.

The investigation of the cognitive impact of live professor presence is a unique and critical first manipulation to the classroom setting. Experiments on the impact of social presence have indicated a cognitive and behavioural impact of feeling as though one is being watched (Bond & Titus, 1983; Foulsham et al., 2010; Huguet, Galvaing, Monteil, & Dumas, 1999) and that the possibility for social interaction impacts behaviour (Aragon, 2010; Laidlaw, Foulsham, Kuhn, & Kingstone, 2011; Richardson, & Swan, 2003). A live classroom setting employs both these social presence factors: the professor being physically present might impact behaviour (e.g., whether a text message is sent, how focus of attention is maintained) and there is also the potential for interaction (e.g., being asked to answer a question). Furthermore, I believe that the instructor presence is perhaps the most critical change between live and video-learning settings. Beyond implications of the effects on human behaviour listed above, the role of professors as instructors is a crucial part of the university objective. Institutions are ranked for their teaching capabilities and many instructors must carefully balance their time between teaching, research and service. Determining whether professor presence impacts the learning experience is an important question for university initiatives. On the one hand, professors are typically required to perform teaching duties, and must consider the value they impart on the teaching experience by being physically present. On the other hand, pre-recording lectures could free up professor resources allowing for more one-on-one time with students.

By manipulating only live professor presence, I maintain all other factors that might vary between a live and video-based lecture setting: the social presence of peers, consistent pacing (e.g., not stopping the task and re-starting at a later time), physical environment (e.g., lecture hall) and any other factors that might differ between settings. This means that any effects found can be attributed to the single manipulation: live professor presence. One additional benefit of the present experiment's design is that it is more ecologically valid than traditional lab-based experiments (Kingstone et al., 2008). That is, the participants are students enrolled in the Introductory Psychology class where the experiment takes place and the materials being presented are lectures in the course that deliver information relevant to the students successfully passing the course. This stands in sharp contrast to testing students individually in a laboratory on information that has little if anything specific to do with any course they are taking, and where motivation for taking part may be supplemental participation credit or some nominal form of remuneration.

3.1.1 Method

3.1.1.1 Participants

An invitation to participate in the experiment was extended to students registered in two sections of Psychology 102: Introduction to Developmental, Social, Personality, and Clinical Psychology at the University of British Columbia. A total of 276 students (180 female; M_{age} = 19.84 years, SD_{age} = 3.41) participated in both sessions of the experiment in return for course credit. An additional 82 participants (Live only = 37; Video only = 45) did not complete both sessions of the experiment and so were excluded from analyses (with one notable exception which I flag for attention in the Results section). In the present experiment power analyses were not conducted as sample was limited to those registered in the class, and those who attended the classroom session on the testing days.

3.1.1.2 Lecture materials

Content was pre-determined by the professor. Two lectures were presented that encompassed the theme of Treatments of Clinical Disorders, but differed in topic: Psychotherapy and Drug Treatment. Session 1 (Live) differed for each class section, such that one class was presented material on Psychotherapy and another received material on Drug Treatment. These lectures were video-recorded by a researcher and then presented for the second session of the experiment (Video version). This allowed me to change the topic for Session 2 (Video) so that students did not see a repeat of the same topic. Each lecture was approximately 60 minutes in length. Lecture material was supplemented with PowerPoint slides.

The order of conditions led to an unbalanced design however there are practical, conceptual, and statistical reasons for. First, a practical consideration: the professor needed to present the material so that I could record it and I did not want to tamper with her presentation style in any way that could negatively impact the ecological validity of the experiment or quality of the Live versus Video comparison (e.g., have her pre-record the lectures in her office or in front of an empty classroom). Therefore, I recorded the Live lectures that came immediately before the Video lecture presentations, thereby minimizing any other differences between conditions, while counterbalancing the lecture topic. A conceptual consideration is that, in actuality, the live lecture condition is just one of many live lecture experiences that the students have had, and will continue to receive, in their educational experiences, and thus its occurrence is not uniquely defined with regard to the Video condition, i.e., as coming before or after the video condition. Finally, I reasoned that if there was any effect of order, placing the Video lecture second would, if anything, operate against the finding of an advantage for the Live lecture relative to the Video lecture, i.e., it represented a conservative test of the working hypothesis that Live lectures enhance student performance. In other words, student performance in the Live lecture would suffer any cost of being first, and the Video lecture any benefit of being second (e.g., the students in the Video condition could expect to be tested on the material that occurred at the time that they were probed for MW). As I report in the results, these intuitions were validated and a Live lecture advantage may be a conservative estimate of the actual difference between Live and Video lectures.

3.1.1.3 MW probes

Over the course of each lecture 6 MW probes were displayed on lecture slides. The timing of these slides was predetermined, such that probes occurred just after the presentation of lecture material that would later be part of the retention test. The slide presented the question "In the moment prior to this slide, were you mind wandering?" (based on Lindquist and McLean, 2011, p. 161) and students were asked to circle "Yes" or "No" on the response sheet provided. Participants were given an average of 24 seconds to record their responses.

3.1.1.4 Retention test

To identify differences in retention between Live and Video lecture settings a 6 item True-False retention test was given at the end of each lecture. Retention test questions were presented on a PowerPoint slide at the end of the lecture and responses were recorded on the participant response sheet. Students were asked to work independently and not use their notes as aid to answer the questions.

3.1.1.5 Ratings

To investigate whether ratings of interest differed between the two learning environments, participants were asked to rate their interest in the lecture topic on a 5 point scale at the end of each session (where 1 = very little interest and 5 = high interest). To investigate whether motivation to remain attentive differed between the two lecture settings a 5 point scale was created and included on the participant response sheet (where 1 = very unmotivated and 5 = very motivated). To further compare motivation across the lectures, at the second session (Video lecture) students were also asked to rate which version they were most motivated to attend to: Live, Video or Equally Motivated.

3.1.1.6 Procedure

Participants completed the experiment during their regularly scheduled class, across two class sessions. Prior to the experiment date a researcher attended the class and explained the broad nature of the experiment to the students and, critically, that they were under no obligation to participate in the experiment, nor would the decision to not participate have any impact on their course grade. The day of each lecture, students were provided separate informed consent and response sheets and given instructions for the experiment by the researcher. The instructor did not provide any information about the experiment to the students. Students were asked to provide basic demographic information on the participant response sheet. Students were provided a definition of MW: "Any thoughts that are experienced that are not related to the material being presented" and examples were provided (e.g., thoughts about lunch; thoughts about past weekend events; concerns about coursework, etc.). There were no differences in instruction provided to the two sessions. Upon completion of the class participants turned in the consent form and response sheet to the researcher.

3.1.2 Results

I counterbalanced across lecture topic as it was not a variable of interest in this experiment. Nevertheless, I conducted a mixed effect model analysis with topic as a factor to examine if it might have a differential impact on the data. This analysis revealed that topic was a significant factor for MW reports, but not for any other variable. There were no meaningful interactions in the data. This analysis allowed me to be confident that collapsing across the two counterbalanced lecture topics to increase statistical power would not compromise the results and interpretations of the data.

Retention test performance, MW reports, motivation and interest ratings were all assessed using paired samples tests, with Lecture Style as the Independent variable. Table 3.1 displays a summary of descriptive statistics.

Measure	Live	Video
Retention performance (Proportion) (Range 0-1)	.74 (.20)	.70 (.21)
MW rates (Proportion) (Range: 0-1)	.48 (.23)	.49 (.27)
Interest rating (Range: 1-5)	3.08 (.82)	2.84 (.87)
Motivation rating (Range: 1-5)	3.63 (1.01)	3.20 (1.16)

Table 3.1. Summary of task means, with standard deviations in parentheses.

3.1.2.1 Content retention test

A paired-samples t-test indicated that participants performed significantly better on the Live lecture retention test compared to the Video lecture retention test, $t(275) = 2.83$, $p = .005$, 95% CIs [.014, .078], $BF_{10} = 2.5$, $d = .17$. Note that in the Methods section it was proposed that the unbalanced design, with Live preceding Video sessions, may actually yield a conservative estimate of the actual difference between Live and Video retention performance. This is because participants in the Video session could use their previous experience in the Live session to correctly anticipate that they would be tested on lecture material that co-occurred with the MW probes. To test this I compared accuracy for the participants who only completed the Video session (n = 45, and as such received the Video session "first") against a randomly selected subset of Video condition participants (n = 45) who attended both the Live and Video sessions (and therefore, received the Video condition second). This analysis revealed that those who received the Video condition second performed significantly better (M = .71) on the retention test compared to those who received the Video condition "first" (M = .61), $t(88) = 2.2$, $p = .03$, 95% CIs [.02, .18], d = .24. This confirmed my hypothesis, that the enhanced retention performance of those in the Live condition, who received the test material first, relative to the retention performance of those participants who received the Video condition second, is a conservative estimate of the benefit of receiving a Live lecture.

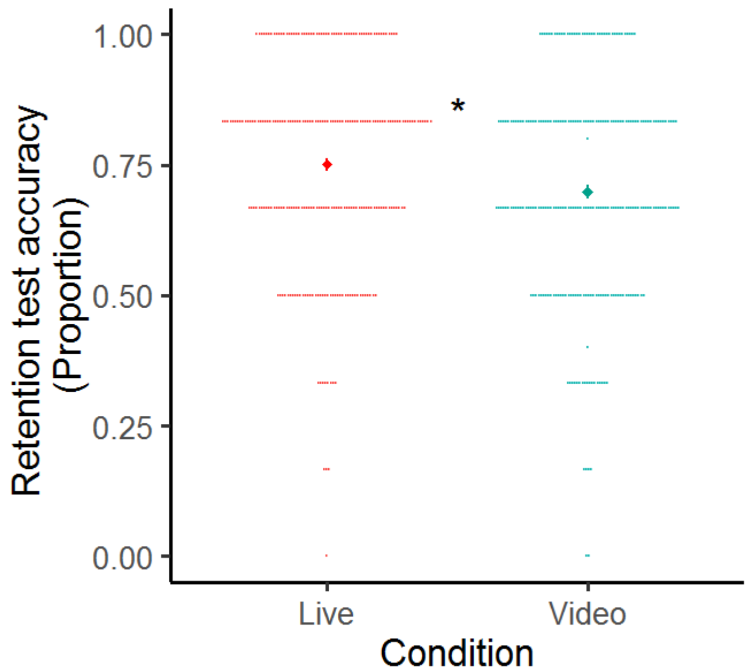


Figure 3.1. Retention test performance between sessions. Bars represent one standard error of the mean.

3.1.2.2 Mind wandering reports

A paired-samples t-test revealed no significant effect of Lecture Style, $t(275) = .81$, $p = .42$, 95% CIs [-.04, .018], $BF_{01} = 15.07$, $d = .05$.

3.1.2.3 Interest and motivation ratings

Twenty-one students did not complete the interest and motivation sections of the response sheet for one or both of the sessions, meaning that their data could not be included for these analyses (N = 255). Interest ratings were significantly higher (indicating greater interest) in the Live lecture compared to the Video lecture, $t(255) = 4.32$, $p < .001$, 95% CIs [.13, .34], $BF_{10} > 150$, d = .27. As per Interest ratings, an ordinal scale was used to measure Motivation rating and so a Wilcoxon signed rank tests was completed. This test revealed that Motivation Ratings were significantly higher (indicating greater motivation) in the Live lecture compared to the Video lecture, $t(255) = 5.28$, $p < .001$, 95% CIs [.26, .57], $BF_{10} > 150$, $d = .33$.

In the second session (Video condition) participants were asked to reflect on both sessions and select which lecture they had been most motivated to attend to. Of the participants who completed this section of the response sheet ($N = 265$), 61% of individuals reported being more motivated to attend the Live lecture, 28% reported equal motivation and only 11% reported they were more motivated to attend to the Video lecture, a chi-squared one-variable test indicated that the ratings were significantly different, X^2 (2, $N = 265$) = 101.65, $p < .001$. These results are consistent with the Motivation Rating reports made at the end of each session which revealed greater motivation to attend to the Live lecture over the Video lecture.

3.1.2.4 Correlational analyses test

Prior work has investigated the impact that interest and motivation have on cognitive factors such as content retention, and MW (Forster & Lavie, 2014; Giambra & Grodsky, 1989; Lindquist & McLean, 2011; Sankaran & Bui, 2001; Unsworth & McMillan, 2013). Unfortunately, most of these experiments did not take place in classroom settings (Lindquist & McLean, 2011; Sankaran & Bui, 2001). As such, I investigated which, if any, of these correlations extend to an actual classroom setting. Correlations can be found in Table 3.2.

	MW - Live	Interest Rating - Live	Motivation Rating - Live	MW - Video	Interest Rating - Video	Motivation Rating - Video
Retention performance - Live	-.08	0.14*	.11			
MW rate - Live		-0.29**	-0.22**			
Retention performance - Video				-0.16*	0.18*	0.17*
MW rate - Video					-0.32**	-0.35**

** $p < .001$, * $p < .05$

Table 3.2. Pearson's correlations for factors of interest (*N* = 265)

The correlation between MW and retention was found to be non-significant for the Live lecture, however was significant for the Video lecture, indicating that increased MW reports were related to decreased retention test performance. Previous research suggests that Interest and Motivation Ratings correlate positively with retention performance, such that greater motivation or interest ratings are related to higher retention performance. All correlations were in this predicted direction, with significant correlations for all but retention and motivation in the Live condition. As per Lindquist and McLean (2011), I predicted a negative relationship between MW and interest, where greater Interest ratings would be related to fewer mind wandering reports. Indeed, correlations were significant for both Live and Video sessions. Finally, and as per research in non-classroom domains, there was a negative relationship between MW and motivation. This result suggests that lower motivation ratings are associated with a higher number of MW reports, for both Live and Video sessions.

To determine whether the relationship between MW and retention in the Video condition ($r = -.16$) was significantly larger than that Live session ($r = -.08$) a comparison of correlation *r*s was conducted. The $z_{\text{Difference}}$ of these scores was .91, $p = .18$, thus suggesting that the difference in the magnitude of the correlations was not significantly different.

3.1.3 Discussion

The current experiment investigated how different learning environments (live in-class versus pre-recorded video) impact two aspects of the learning experience: retention performance and MW. I also examined how subjective experiences, specifically interest and motivation, differ between these environments and what impact, if any, they have on retention and MW. The finding that retention performance was significantly higher in the Live session compared to the Video session suggests that the acquisition and retention of the lecture information benefitted from having the lecture delivered by a professor who was physically present in the classroom. Moreover, as noted previously, the design of the present experiment was such that this significant difference is, if anything, an underestimation of the actual magnitude of the Video lecture disadvantage. Perhaps the most obvious and important implication of this finding is that in real world settings, where lecture material is delivered in an online format using lengthy video content (e.g., MOOCs), students taking such courses may retain much less lecture information than those who receive the material in a classroom with a professor.

Interestingly, the difference in retention performance between Live and Video instruction cannot be attributed to differences in attention vis-à-vis MW as the results indicate that MW was unaffected by the lecture format. That said, the correlation data reveal that retention declined as MW increased in the Video condition, but a similar relationship was not observed in the Live condition (see Table 3.2). The difference in the magnitude of these correlations was not significant but the possibility that retention for video-based material is more sensitive to shifts in attention could be a question worthy of future research. As discussed in detail below, the relationship between MW and retention performance in the video session could be because the material is found to be less interesting, and the students are less motivated, when watching a video lecture than when they receive it live.

The change in interest for the material is most critical, as the correlation data indicate that for both lecture types retention performance changes with a shift in interest, i.e., retention performance declines when interest declines. It is my position that this effect is a conservative estimate of the real-world impact of delivering course material online via video-based lectures to students. The reason for this being that a number of factors were present in this design that would not be in a pure online setting, for example: (i) it was a novel experience, (ii) the students received the video information in their usual classroom and normal lecture time, and (iii) students were surrounded by their peers, which offers a well-established social benefit (Erichsen & Bolliger, 2010; Garrison, 2011; Sung & Mayer, 2012). Thus my experiment suggests that not only may students' retention of course material suffer relative to their in-class cohorts who receive live lectures, but with the decline in interest and motivation, students may well struggle to keep up with a course; a possibility that converges with the high attrition rates for MOOCs: typically over 90% of the students enrolling in MOOCs fail to complete them (Jordan, 2015; Khalil & Ebner, 2014; Meyer, 2012; Onah, Sinclair, & Boyatt, 2014; Pomerantz, 2014).

The present experiment spearheads a much needed initiative for careful examination of differences in cognitive impact and subjective experience between traditional classroom learning and video-based online learning environments. I recognize that there may be benefits to video learning (e.g., material can be re-watched, students can watch at a more convenient time, etc.) or at least different ways that students can interact with materials compared to with a live lecture (Demetriadis & Pombortsis, 2007; McKinney et al., 2009; Wuensch et al., 2008). Whether these benefits outweigh the costs I have reported here is an important issue for future investigation. With growing demand for flexible and convenient online video-based learning settings, institutions and instructors should be mindful of the cost of simply moving content to an entirely online video-based method, without making amends for cognitive impacts could mean compromising a tried-and-true learning experience. This research is especially important in the context of flipped-classroom models, which often encourage video-based learning outside the classroom to facilitate increased in-class discussions. The impact of these video-based activities should be examined carefully with respect to how MW and content retention may be impacted. A related consideration is that video-content length has been examined with regard to an “ideal” video length, and many educational researchers suggest that shorter is better (Bradbury, 2016; Crook & Schofield, 2017; Schacter & Szpunar, 2015). Thus, an examination of whether the effects found in the present experiment generalize to shorter video content would be of great value.

In sum, the present experiment illustrates that students' retention performances differ between live and video versions of a lecture, and that they are more motivated and interested when attending a session with a live professor present. These results further indicate that natural settings can be used to investigate the differences in learning behaviours and experiences, and that comparisons need not be limited to laboratory environments. Given that a single change to the classroom environment (physical professor presence) impacted cognitive processes, future experimental paradigms could help to tease apart the impact of other factors that are more difficult to manipulate in naturalistic settings. The next lab-based experiment extends this important work through the manipulation of additional factors (i.e., peer presence, instructor visibility) that differ between live and video-based lecture settings.

3.2 Experiment 7: The impact of learning environment on mind wandering and material retention

As indicated in Experiment 6, online video-based lectures vary from live lectures in many ways, such as social context, peer presence, regularity of meetings, ability to interact with content, etc. However, even among online lectures there is little consistency in presentation format, with some courses using only slides and voice-over recordings whereas others may include a video-recording from an actual lecture. In terms of social presence, students may participate in online lectures in a variety of settings: alone in their own space, in a shared setting (e.g., library) or as part of a satellite course with other students. Given the diversity of online lecture dynamics, I wanted to examine two features that often vary when taking an online course: the presentation of the lecture (professor visibility) and the learning setting (presence of others).

The past decade has witnessed online learning environments become a staple in both university settings and in open forums (e.g., edX, Coursera, etc.). In the United States and Canada, more than 22% of college students report taking at least one online course (Abdous & Yoshimura, 2010; Canadian Digital Research Learning Association, 2019). Although significant efforts have been devoted to modelling and understanding student engagement and learning during online courses (Srećko Joksimović et al., 2018), there is relatively little experimental work exploring how learning and engagement are affected by the way the online lectures are presented or where the student watches the lectures. This gap in the literature is particularly concerning given how pervasive MW is in educational settings (Szpunar, Moulton, et al., 2013). The prevalence of MW, combined with its frequent negative relationship with comprehension, highlights the importance of understanding how features of online learning influence MW. The current experiment addresses two potentially critical features: 1) the importance of being able to see the instructor during an online lecture and 2) the social setting within which the lecture is viewed.

Online lectures provide an opportunity to easily change and compare features of a learning environment. By making small changes to the display or content, researchers can gain a better understanding of what keeps students engaged. Indeed, student engagement in online lectures is an understudied but critical factor, given the high rates of online lecture drop out (Jordan, 2015; Khalil & Ebner, 2014; J. Kim et al., 2014) and increasing MW rates over long intervals (Szpunar, Khan, et al., 2013). The high drop-out rates, however, do not necessarily outweigh the benefits of reaching more learners with distance education and the ability to retain valuable learning analytics information that can be used to improve learning over time.

As noted earlier in this chapter, previous work on the effectiveness of pre-recorded lectures has yielded inconsistent results with respect to their effectiveness compared to live lectures (Bernard et al., 2004). For example, Sankaran and Bui (2001) found that the use of deep or surface learning strategies did not differentially influence motivation or performance between web and live lecture settings. In contrast, McKinney et al. (2009) found that students who listened to a lecture recording via podcast performed better on an exam than those who attended a live lecture. More recently, Varao-Sousa and Kingstone (2015) found that material retention was worse in pre-recorded lectures versus live lectures, but MW rates were not impacted. Given the inconsistent findings with regard to the effectiveness of pre-recorded video lectures, it is plausible that features either within the video lectures or within the student's environment may be impacting engagement and effectiveness. Such features are the focus of the current experiment.

A few experiments have capitalized on the idea that features of the lecture may impact the learning experience by manipulating specific features in online lectures, such as length of lecture content, and video display content (Kizilcec, Papadopoulos, & Sritanyaratana, 2014; Wilson et al., 2018). One feature of online learning that is often inconsistent in online learning displays is whether an instructor is actually visible. The idea that instructor visibility may influence attention is well motivated by experiments indicating that socially-relevant or sociallyrelated information (i.e., faces) captures our attention (Langton, Watt, & Bruce, 2000; Theeuwes & Van der Stigchel, 2006). Thus, the visibility of an instructor could impact the way that students actively attend to material (for example, by modulating MW reports).

Chen and Wu (2015) tested this idea by manipulating lecture format, with students watching the lecture in one of three possible ways: a pre-recorded version of a lecture, a visual of lecture slides with the instructor voicing over the content, or a visual of lecture slides with a video of the instructor inset on the screen. Their results indicated that sustained attention (inferred via a single-electrode EEG system) was better in the voice-over compared to the instructor overlay format, yet learning performance was the worst in the voice-over condition compared to the other two. A more recent investigation manipulated lecture presentation using three conditions: audio-only, audio with text, and audio with instructor visible (Wilson et al., 2018). In this experiment, neither MW nor comprehension were affected by changes in content display. The mixed results found across these two experiments, with respect to the impact of instructor visibility on attentional processes, may stem from the different stimulus manipulations and MW measurements. In the present experiment, I sought to provide a specific manipulation by varying only the instructor's visibility in the video frame, and to test for its effect on MW and material retention.

Aside from visibility of the instructor, the learning setting of online material can be highly variable. As noted above, cues related to social relevance can influence attention capture, and online lecture watching can differ in terms of social presence. For example, students may view content alone in their own space, in a shared setting (e.g., library), or as part of a satellite course with other students. Thus, the mere presence of others or the solitary nature of a setting could influence student attention to material, and by extension, material retention.

Abdous and Yoshimura (2010) examined if final grades were influenced by whether a course was taken in a live class, through a video broadcasting site, or via live video-streaming. Their results suggested that setting did not impact grade or course satisfaction. However, this study did not collect any measurement of attention and the lectures were all presented in real-time. Given that many students in online courses watch pre-recorded videos in a multitude of settings, the current experiment sought to address if social setting mattered while viewing pre-recorded videos. I also extend previous work by examining shared spaces as a social setting for online learning. This setting differs from work examining differences between live and online lecture formats (Bernard et al., 2004; McKinney et al., 2009; Sankaran & Bui, 2001; Varao-Sousa & Kingstone, 2015), as students are often in group settings without a shared, scheduled event (e.g., on campus, libraries, coffee shops).

To better understand the impact of diversity in online lecture presentation and learners' environments, I experimentally manipulated two features that often vary when taking an online course: 1) the presentation of the lecture (instructor visibility) and 2) the learning setting (the presence, or lack, of others). The following research questions are addressed: What are the main effects of presentation and setting on MW, and on comprehension? What is the relationship between MW and comprehension, and does social setting or instructor visibility moderate this relationship?

3.2.1 Method

A 3x2 between-subjects design was employed. Social setting was manipulated in three ways, with students completing the experiment either: alone in a laboratory testing room (Alone), in a classroom with other students collectively viewing a single projector screen (Group – shared screen), or in group with other students but each student viewed the lecture on their own computer monitor (Group – individual screens). The visibility of the instructor was manipulated in two ways: either the instructor was fully visible in the video recording (Instructor Present), or the instructor was cropped out of the video recording (Instructor Absent).

3.2.1.1 Participants

A power analysis for a 3x2 between-subject ANOVA was conducted using G*power (Faul et al., 2007). A predicted medium effect size of .25 (based on effect sizes reported in Tu & McIsaac, 2002: ANOVA comparison of social context ratings in classroom discussion forums; ANOVA comparison of attention when manipulating lecturer size and visibility: Korving, Hernandez, & De Groot, 2016), .80 power, and a set alpha of .05, indicated that 244 participants would be required. Due to running subjects in groups, data was collected for an additional two weeks to ensure that sample size requirements would be achieved. Two hundred and seventy-nine undergraduate student participants (M_{age} = 20 years, SD_{age} = 2.1 years, 198 women and 81 men) were recruited from the University of British Columbia and participated for course credit. Ethics approval was obtained from UBC's Behavioural Research Ethics Board. Written, informed consent was obtained from each student at the start of the experiment.

3.2.1.2 Lecture materials

A 58 minute video recorded lecture on Psychotherapy was used for all groups (adapted from Varao-Sousa & Kingstone, 2015). For the instructor-absent condition the video display was cropped so that only the lecture slides (and not the instructor) were visible.

3.2.1.3 MW probes

Across the lecture six MW probes were displayed visually on the lecture slides at the following time points: 4:30, 13:45, 20:02, 31:52, 44:27, and 57:20 minutes. These probes were embedded within the lecture content used, and originally selected to appear following every 3-4 slides. The probes asked: "In the moment prior to this slide, were you mind wandering?" and students responded by circling "Yes" or "No" on the response sheet provided. Students had an average of 30 seconds to record their responses before the lecture continued.

3.2.1.4 Retention test

A six-item True-False test based on the lecture content was administered via slideshow at the end of the lecture. Students recorded their responses on the participant response sheet.

3.2.1.5 Other measures

I also counted the number of note pages students made during the lecture and assessed interested and motivation using single questions. A 5-point Likert Scale was used, where 1 = Low Interest/Motivation and 5 = High Interest/Motivation. There were no differences across conditions, thus alleviating any concerns of confounding effects of notes, interest, or motivation between conditions. I do not discuss these measures further given the scope of the paper.

3.2.1.6 Procedure

An online booking system was used to recruit students for the experiment. For each of the three experimental conditions, a research assistant explained the nature of the experiment to the students, and provided separate informed consent, response sheets and lined paper. The following definition of MW was provided: "Any thought that is not related to the lecture material being presented" and examples were provided (e.g., thoughts relating to getting lunch later, thinking about something that happened on the weekend, etc.). Students were asked to respond on the participant sheet, and told that they were free to make notes during the video as they would during an actual class. Students were also informed that they would complete a short comprehension test at the end of the session, during which time they would not have access to their notes. Students in the group-individual screens condition wore headphones while viewing the lecture. The video was started by the researcher, and then the researcher exited the room, leaving the student(s) alone to view the lecture. At the end of the video, the student/s submitted any notes that were taken to the researcher and then completed the retention test, as well as indicating their interest and motivation for the material in the video.

3.2.2 Results

Table 3.3 presents the descriptive statistics for the variables measured across each level of design. I discuss the results for each of my research questions below.

	Instructor Visible			**Instructor Absent**		
	Alone	Group - Shared Screen	Group - Individual Screens	Alone	Group - Shared Screen	Group - Individual Screens
Main Measures						
MW (Range 0-1)	.45 (.22)	.57 (.25)	.50 (.25)	.47 (.22)	.52 (.27)	.48 (.29)
Retention (Range 0-1)	.80 (.15)	.76 (.21)	.74 (.19)	.82 (.17)	.77 (.22)	.80 (.21)

Table 3.3. Mean reports of dependent measures, with standard deviations in parentheses.

3.2.2.1 Mind wandering reports

There was no evidence for a main effect of instructor visibility on MW, $F(1, 273) = .46$, $p = .50$, $BF_{01} = 6.29$, $\eta p^2 = .002$ and no interaction between instructor visibility and social setting, $F(2, 273) = .52$, $p = .50$, $BF_{01} = 12.67$, $\eta p^2 = .004$). However, I did find a marginal evidence for an effect of social setting on MW, $F(2, 273) = 2.79$, $p = .06$, $BF_{10} = .48$, $\eta p^2 = .02$. Post-hoc comparisons (Bonferroni correction) of social group indicate that students who watched in the group – shared screen condition reported MW marginally more than students who watched the lecture alone in the laboratory, $t(92) = 2.44$, $p = .02$, 95% CIs [-.17, .003], $BF_{10} = 1.96$, $d = .36$ (Figure 3.2). In contrast, MW reports did not differ between the group-individual screens and the other two conditions (group-shared screen and alone), $t(92) = 1.37$, $p = .17$, 95% CIs [-.04, .14], $BF_{01} = .51$, $d = .20$ and $t(92) = .92$, $p = .36$, 95% CIs [-.12, .06], $BF_{01} = 5.80$, $d = .13$, respectively.

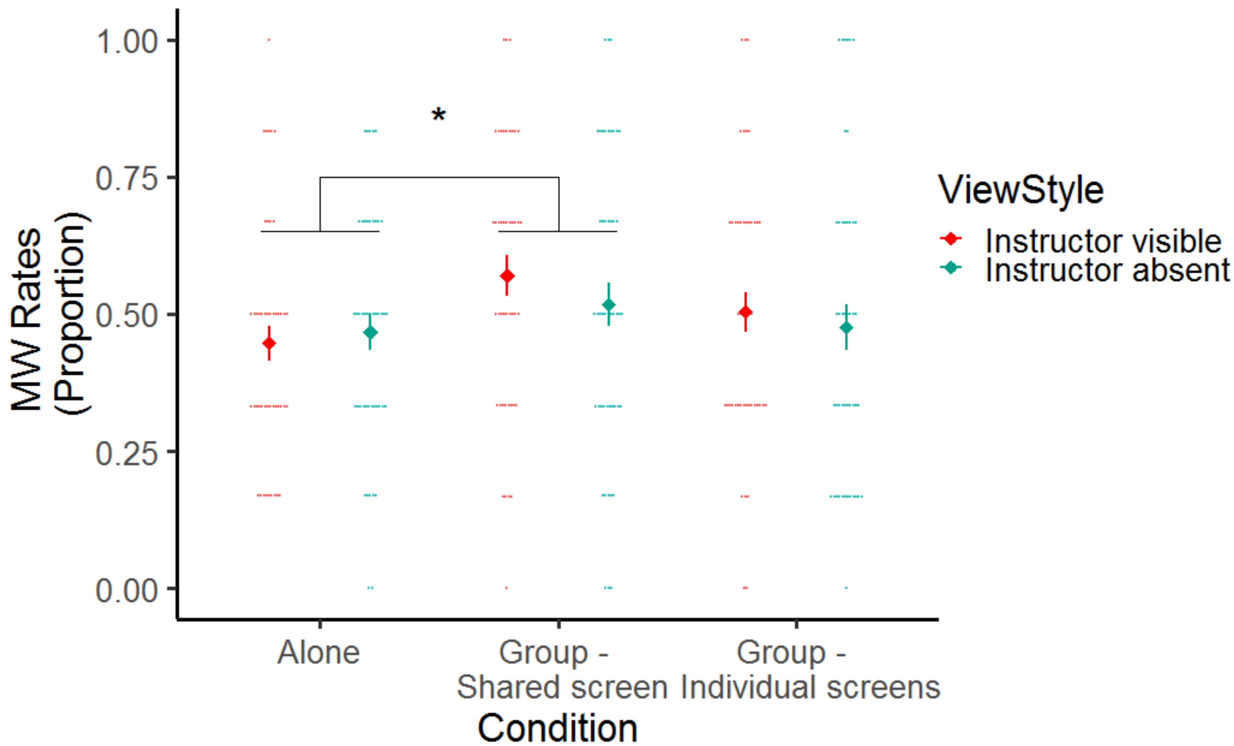


Figure 3.2. Probe-caught mind wandering reports. Bars represent one standard error of the mean.

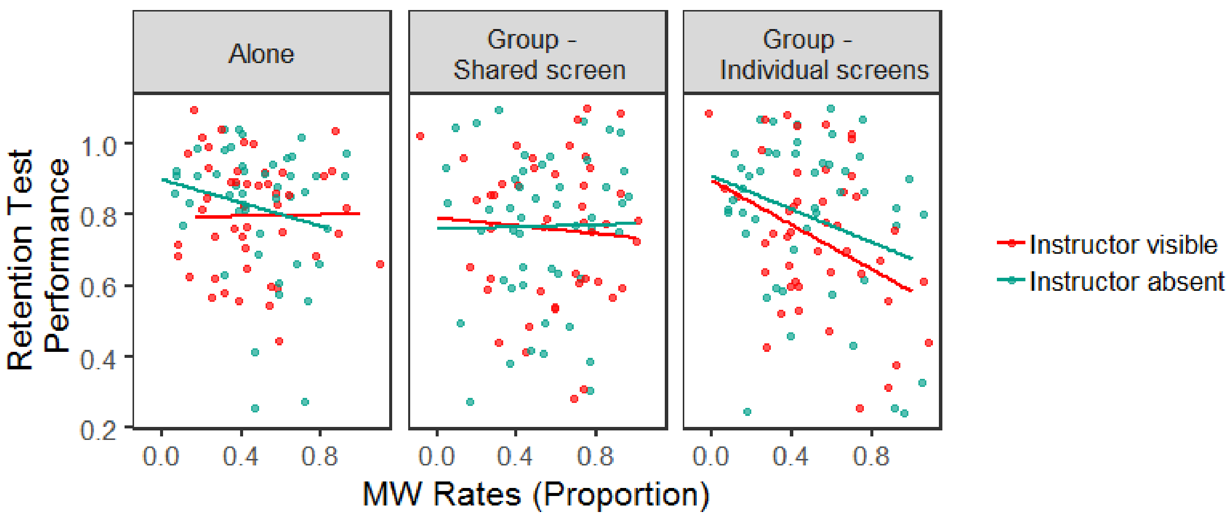
3.2.2.2 Retention test

There was no significant main effects of social group or instructor visibility and no interaction, $F(2, 273) = 1.61$, $p = .20$, $BF_{01} = 3.31$, $\eta p^2 = .01$; $F(1, 273) = 1.83$, $p = .18$, $BF_{01} = 6.10$, $\eta p^2 =$

.007) and $F(2, 273) = .42$, $p = .66$, $BF_{01} = .003$, $\eta p^2 = 19.61$, respectively. Thus, all students performed similarly well, regardless of what version of the lecture they saw or in which setting they watched the video.

3.2.2.3 Mind wandering and retention relationship

I next examined the relationship between MW and retention, and whether the instructor visibility or social setting moderated this relationship. First, I regressed average retention scores on average MW separately to assess the overall relationship. As expected, retention was significantly, and negatively related to MW, $B = -.14$, $SE = .04$, $p = .003$. I also tested if social setting or instructor visibility moderated the relationship between MW and retention – i.e., is it



possible that MW shared a more negative association with retention in certain conditions? I tested this by computing two-way interactions: 1) MW × social setting and 2) MW × presentation style, see Figure 3.3.[5]

[5] We did not have enough power to test for a three-way interaction (requiring roughly four-fold the sample size needed for a two-way interaction, Heo and Leon (2010)).

Figure 3.3. Correlation plots of relationship between mind wandering rates and retention test performance, split by group.

There was no evidence that instructor visibility moderated the relationship between MW and retention. However, there was a significant two-way interaction between MW and social setting, $F(2, 273) = 3.22$, $p = .04$, $\eta_p^2 = .02$. MW had a stronger negative relationship with retention for students in the group-individual screens condition, $B = -.27$, $SE = .07$, $p < .001$, compared to group-shared screen ($B = -.02$, $SE = .08$, $p = .83$) or alone conditions ($B = -.08$, $SE = .08$, $p = .34$).

3.2.3 Discussion

The present experiment investigated whether MW and/or retention would be influenced by instructor visibility or social setting. The results indicate that instructor visibility does not impact MW or retention, but there is evidence to suggest that social setting may influence MW rates. MW was higher in the group – shared screen condition compared to watching the lecture alone. These results also replicated previous experiments suggesting that MW and retention are negatively related (Dixon & Li, 2013; Smallwood, McSpadden, et al., 2007; Varao Sousa et al., 2013), and extend these findings by showing that certain social conditions may amplify this effect. In particular, this negative relationship was the strongest when students were in a group setting but viewing the lecture on their own individual screens.

Although results suggested that the visibility of the professor did not influence the results, future work may wish to investigate other online course display methods (e.g., audio without visual; text without audio or visual, etc.). Work by Wilson et al., (2018) compared multiple online viewing displays and reported that visibility of the instructor actually impaired comprehension although participants reported a preference for lectures with the professor visible.

Research by Wammes and Smilek (2017) found that MW rates increased over the course of the lecture for online viewers, in contrast to live viewers. These two highly related experiments suggest that there are factors which influence MW rates and online lecture preferences, and that experimental contrasts between online formats and live lectures should be further explored.

Future lines of work could be developed to determine why group setting has an impact on MW rates. Research suggests that the ability to see a peer's gaze direction or activity during learning can influence one's own attention (Phillips, Ralph, Carriere, & Smilek, 2016). Although there was no collected evidence in the present experiment that students in the group-shared screen condition had more peer distractions, prior work does suggest that individuals' attention is strongly biased to, and by, the presence of others (Langton et al., 2000; Theeuwes & Van der Stigchel, 2006). Although researchers have examined the benefits of peer engagement in online courses (d'Alessio et al., 2019; S. Joksimović, Gašević, Kovanović, Riecke, & Hatala, 2015; Richardson, Maeda, Lv, & Caskurlu, 2017), determining the specific ways in which the physical presence of others influences student success would benefit research related to promoting inclusion in online courses. Future work could also examine why there was a negative correlation between MW and retention in the group-individual screens condition, particularly since this condition did not display the highest overall rates of MW. One possibility is that participants experienced different types of inattention that were more harmful to learning, such as distraction—a distinct form of inattention (Unsworth & McMillan, 2014). For example, students may have been distracted by the mere presence of their peers and the inability to monitor what others were doing. As the present experiment only measured rates of MW, it is possible that distractions may have also influenced the underlying relationships found. A better understanding of factors that influence learners' cognitive states during online settings will clearly be valuable to predictive models of in attention and learning.

This experiment is not without limitations. First, despite testing nearly 300 students some of the effects may have been underpowered for detecting small effect sizes: For example, a sample size of over 320 is recommended to detect a 'small' interaction effect (d = .20) with a power of .90. Second, although I attempted to mimic three ecological settings in which students watch online lectures, devising ways to truly capture students in their naturalistic settings can only enhance the external validity of the findings.

The steady increase of students partaking in online learning material is contrasted with a lack of understanding about optimal features of the lecture and one's own environment. I provide some insight to this issue and suggest that future work may benefit from considering more contextual features such as social setting. Specifically, knowing where a student is while viewing the lecture may inform predictive models of student attention and engagement.

3.3 Experiment 8: MW in a leisure environment

Daily experience sampling outside of the lab suggests that MW occurs between 30-50% of our lives across activities ranging from reading, driving, playing sports, and while conversing (Kane, Gross, Chun, Smeekens, Meier, Silvia, Kwapil, et al., 2017; Killingsworth & Gilbert, 2010; McVay, Kane, & Kwapil, 2009). Despite how frequently MW is reported in everyday activities, most research of the cognitive state investigates rates in simple, unnaturalistic laboratory-based tasks (e.g., the SART, attentional blink paradigms, and word-by-word "reading" tasks). These tasks are often chosen as their simple designs can hone in to specific mechanisms of attention, but moreover, they are known to be dull, unexciting tasks that will elicit mind wandering (Hawkins, Mittner, Boekel, & Heathcote, 2015; Jackson & Balota, 2012; Morrison, Goolsarran, Rogers, & Jha, 2014; Seli, Carriere, et al., 2015; Schooler, personal communication).Yet, one could argue that precisely because these tasks demand that people perform boring tasks for a sustained period of time, they are not representative of the way most people spend their free time. Furthermore, there is evidence that boredom influences MW rates, such that boredom occurs when attention cannot be engaged in a satisfying way, resulting in MW being positively correlated with reports of boredom (Eastwood, Frischen, Fenske, & Smilek, 2012; Kane et al., 2007b; Struk, Carriere, Cheyne, & Danckert, 2017). Given that most MW experiments use tasks that are not engaging, I decided to use a task that is extremely engaging, with the goal of determining the lower bound to MW rates. That is to say, perhaps the rates of MW reported both in and out of the lab represent atypically high rates of inattention, as lab-based tasks are unengaging and most measures of MW in everyday settings are conducted in undergraduate lectures. Thus, the goal of the current experiment is to investigate the potential lower-bound of mind wandering rates by using a more naturalistic and engaging leisure activity: playing a video game.

The billion dollar video gaming industry has maintained popularity for decades, appealing to children and adults alike (Entertainment software association, 2017; “Video Game Industry", 2010). Video game popularity is supported by the fact that, as of 2015, roughly 65% of North American households had at least one gaming system and gamers report spending an average of 12 hours a week on video game play, with gamers spending more time gaming than watching TV or movies (Entertainment software association, 2015). Video games are complex activities demand hand-eye co-ordination, attention to detail and speedy processing of constantly changing information and are thus the perfect candidate for an engaging real-world task (Chisholm & Kingstone, 2015).

In addition to measuring MW rates in the natural task of video game play, I examined whether MW rates were influenced by the size of screen used for gameplay as this would potentially provide some variability in response rates and because access to presentation mode varies based on household and player preference. Thus, I examined three formats that individuals may use when gaming: a large projector screen, a computer monitor, and a handheld gaming device. The common media slogan that "bigger is better" has been supported by prior research on attention and arousal to screen size, with research suggesting that increases in screen size positively correlate with levels of attention while viewing images (Bracken, Pettey, Guha, & Rubenking, 2010; De Cesarei & Codispoti, 2006; Reeves, Lang, Kim, & Tatar, 1999; Rigby, Brumby, Cox, & Gould, 2016; cf Bellman, Schweda, & Varan, 2009: no effect of screen size). Therefore, the prediction in the present experiment is that MW rates should decline as the screen size increases.

The goal of the present experiment is to investigate the impact of screen size on MW reports during the leisure activity of video game play. The increased popularity in portable gaming devices in the last two decades afforded this experiment the ability to naturally manipulate and examine the influence of screen size during game play (Ham, 2004; Ivory & Magee, 2009). To accomplish this goal, participants responded to MW prompts while playing video games on three different screen sizes: handheld device, computer monitor, and wall-sized projector. Although attention seems to reliably be held better during passive viewing of larger screens, the results are less clear for active tasks like gaming. Based on the results found for passive tasks, I predicted that MW rates would be lowest when playing a video game in the projector screen condition as compared to the computer monitor and handheld device conditions.

3.3.1 Method

3.3.1.1 Participants

Participants were 36 undergraduate students (M_{age} = 21.5, SD_{age} = 3.9 years, 30 women and 6 men) from the University of British Columbia who participated for course credit. A power analysis for a within ANOVA was conducted using G*power (Faul, Erdfelder, Lang, & Buchner, 2007). A predicted medium effect size (based on an ANCOVA analysis that reported a .45 effect size for immersion tendency across screen size as reported in Hou et al., 2012), with moderate power-to-detect effects (.80), and a set alpha of .05, indicated that 28 participants would be required. In order to conduct a complete counterbalance of the conditions, a total of 36 participants was collected. Participants were pre-screened to select only those with normal or corrected-to-normal vision.

3.3.1.2 Visual platforms

Three screen sizes were used: a projector, a computer monitor and a handheld device. The projector presented the game on a 118 x 154-inch screen at roughly 145” distance (44.29° viewing angle). In the computer monitor condition, the video games were presented on a 21×12-inch screen, with the participant seated roughly 45” from the screen (15.18° viewing angle). For both the projector and computer monitor sessions, video games were played via a Nintendo Wii console, with a wired controller. In the handheld device condition, video games were presented on a Nintendo 2DS console with a 3.53-inch screen which participants held roughly 16” from the face (11.86° viewing angle). The difference in viewing angles was considered acceptable based on (Hou, Nam, Peng, & Lee, 2012) who adeptly suggest that differ sized screens are unlikely viewed at the same distance and accounting for this creates a more natural (albeit less controlled) setting.

3.3.1.3 Video games

Three video games that were compatible with both the Nintendo Wii console and the Nintendo 2DS handheld device were chosen for the experiment. The following games were selected based on a list of best-selling video games: *The Legend of Zelda: the Ocarina of Time* (an adventure-type game), *Super Smash Brothers* (a fighting-type game), and *Mario Kart* (a racing-type game)[6]. Participants played one unique video game for 15 minutes in each of three separate screen sessions.

3.3.1.4 MW probes

During each 15-minute video game session, participants were presented with twelve audio tone probes. Each probe was randomly generated to be spaced at least 45 seconds apart, but with no probe occurring more than 90 seconds apart. The participants were instructed to verbally respond whether they were thinking about task-related thoughts (in this case they would say "game"), or task-unrelated thoughts (in this case they would say "mind wandering") in response to hearing each sound probe. Responses were audio recorded via a laptop situated near the participant.

3.3.1.5 Post-gaming measures

Following each 15-minute gaming session, participants completed an online questionnaire assessing task load, interest and attention. Task load was assessed via the NASA Task Load Index (TLX), which measures subjective workload on a 21-point scale (Cao, Chintamani, Pandya, & Ellis, 2009), I amended this scale to investigate only mental demand and physical demand (where high scores indicated greater demand). Interest was rated on a 5-point scale, where a rating of 1 indicated "no interest", and a rating a 5 indicated "very interested". Finally, participants reported the extent to which the game held their attention, where a rating of 1 indicated "not at all" and a rating a 5 indicated "a lot".

3.3.1.6 Post-experiment questionnaire

At the end of all 3 gaming sessions, participants were asked to report game play frequency (7-point scale ranging from Never – Every day), game preferred (forced choice between the three games), and format preferred (forced choice between the three visual platforms). Following this, participants completed a brief demographics questionnaire.

3.3.1.7 Procedure

After providing informed consent, a participant was brought into the testing room by the experimenter. The experiment consisted of three separate 15-minute video game sessions under each of the screen conditions, with screen size and gaming order randomized across participants. At the beginning of the experiment, the experimenter introduced the overall task to the participant by stating: "Since this is a study on leisure, we will not be evaluating your performance in the game. Your only goal is to have fun, as if you were playing in your living room at home." After this, a sample sound probe was played and participants were informed to respond loudly and clearly in response to each probe. Participants were taught the controls for the video game before each session began. The experimenter returned after 15 minutes, stopped the video game session and administered the post-gaming measures. This procedure was repeated for each of the three sessions. After the last session participants completed the post-experiment questionnaire, and was then debriefed by the experimenter.

3.3.2 Results

A one-way within-subjects ANOVA was conducted on the main measure of interest (mind wandering) to determine whether game played influenced results. Results indicated that game played had a significant influence on the reported outcomes $F(2, 69) = 5.38$, $p = .007$, $BF_{10} = 1.42$, $\eta p^2 = .14$. As game had an influence on the main measure of interest, the analyses for all factors were conducted while controlling for the influence of game played by using an analysis of covariance which included game played in the model.

Table 3.4 provides descriptive statistics for all measures of the experiment. All data were analyzed via a one-way within subjects ANOVA.

	Visual Platform Condition		
Variable	Projector Screen	Computer Monitor	Handheld Device
MW Proportions (Range 0-.67)	.21 (.19)	.22 (.19)	.22 (.20)
Mental Demand (Range 1-19)	9.81 (5.28)	8.69 (4.92)	9.17 (5.66)
Physical Demand (Range 1-19)	4.25 (5.07)	4.00 (4.69)	3.91 (4.98)
Interest Rating (Range 1-5)	3.53 (1.44)	3.53 (1.28)	3.75 (1.25)
Attention Rating (Range 1-5)	3.81 (1.26)	3.75 (1.13)	3.86 (1.13)

Table 3.4. Mean reports of dependent measures, with standard deviations in parentheses.

3.3.2.1 Mind wandering reports

There were no statistically significant differences in MW rates between the three different screen size conditions, $F(2, 70) = .037$, $p = .96$, $BF_{01} = 11.33$, $\eta p^2 = .001$. There was a marginal game by screen size interaction, $F(4, 64) = 2.53$, $p = .049$, $BF_{10} = .015$, $\eta p^2 = .14$, specifically MW was highest when individuals played Super Mario Smash brothers on the handheld device ($M = .27$) as compared to playing Zelda on the handheld device ($M = .13$).

3.3.2.2 Mind wandering time course

A linear trend analysis was conducted to examine whether mind wandering rates changed over the course the task (Figure 1). Previous research has found that MW often increases linearly as time on task increases. Mind wandering increased significantly for the 2DS gaming session, $F(1, 33) = 14.81$, $p = .001$, $\eta p^2 = .31$ and the monitor session: $F(1, 32) = 6.93$, $p = .013$, $\eta p^2 = .18$; but only marginally for the projector session: $F(1, 28) = 3.95$, $p = .06$, $\eta p^2 = .11$. The trend for probe reports of MW to increase over time has been found in numerous experiments within lecture settings (Farley et al., 2013; Lindquist & McLean, 2011; Seli, Wammes, et al., 2015; Varao-Sousa & Kingstone, 2018).

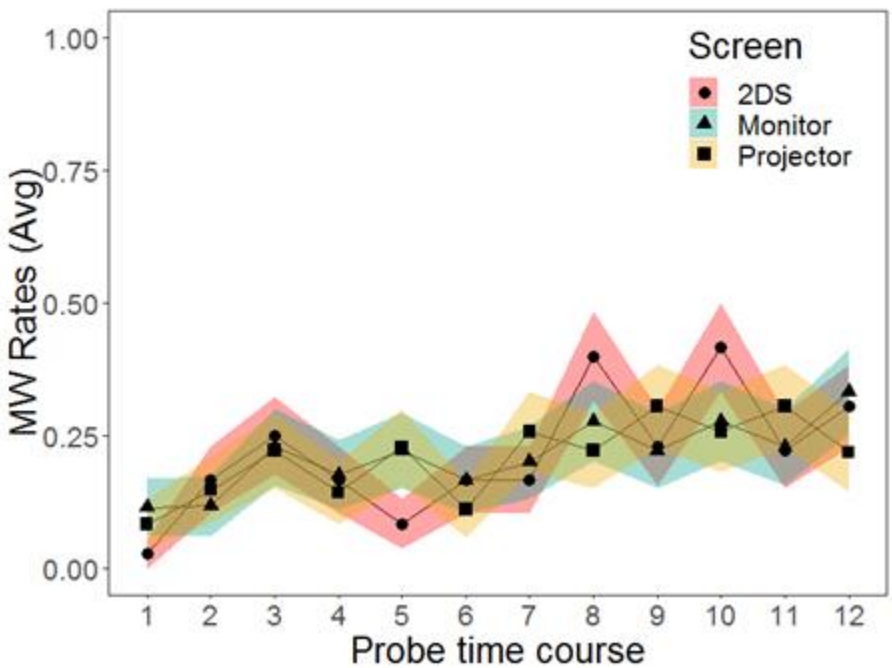


Figure 3.4. Time course of mind wandering reports split by screen type. Coloured bands represent one standard error of the mean.

3.3.2.3 Post-gaming measures

The screen size condition had no statistically significant influence in self-reported mental demand, $F(2, 70) = .68$, $p = .51$, $BF_{01} = 8.39$, $\eta p^2 = .02$, or physical demand, $F(2, 70) = .30$, $p = .74$, $BF_{01} = 11.18$, $\eta p^2 = .009$. Similarly, no differences were found between conditions for interest ratings, $F(2, 70) = .42$, $p = .66$, $BF_{01} = 8.82$, $\eta p^2 = .01$, or attention ratings, $F(2, 70) = .11$, $p = .89$, $BF_{01} = 10.86$, $\eta p^2 = .003$.

When considering the influence of game on the post-gaming measures, game played had no influence in self-reported mental demand, $F(2, 64) = 2.15$, $p = .13$, $BF_{01} = 6.30$, $\eta p^2 = .06$, nor was there an interaction, $F(4, 64) = .10$, $p = .98$, $BF_{01} = 49.26$, $\eta p^2 = .006$. Game played had no influence in self-reported physical demand, $F(2, 64) = 1.13$, $p = .33$, $BF_{01} = 10.79$, $\eta p^2 = .034$ nor was there an interaction, $F(4, 64) = 1.60$, $p = .19$, $BF_{01} > 150$, $\eta p^2 = .09$. Game played did have a significant influence on interest rating, $F(2, 64) = 5.37$, $p = .007$, $BF_{10} = .64$, $\eta p^2 = .14$ and furthermore a significant interaction ($F(4, 64) = 5.87$, $p < .001$, $BF_{10} = .17$, $\eta p^2 = .27$). The interaction revealed that playing Zelda on the projector was significantly more interesting than playing Super Mario Kart on either the TV screen ($t(10) = 3.13$, $p = .011$, 95% CI [.37, 2.17], $d = .98$) or the projector ($t(10) = 3.21$, $p = .009$, 95% CI [.51, 2.77], $d = 1.02$). Similarly, there was an effect of game on attention ratings, $F(2, 64) = 6.54$, $p = .002$, $BF_{10} = 1.24$, $\eta p^2 = .17$, with Zelda ($M = 4.19$) reported as holding significantly more attention than Super Mario Kart ($M = 3.54$) but no interaction between game and screen size, $F(4, 64) = 2.56$, $p = .05$, $BF_{01} = 1.25$, $\eta p^2 = .14$.

3.3.2.4 Post-experiment questionnaire

Sixty-one percent of participants reported enjoying the large projector screen format the most, 19.4% reported enjoying the TV format the most, and 19.4% reported enjoying the 2DS format the most. A chi-squared one-variable test indicated that these ratings were significantly different, $X^2(2, N = 36) = 12.49, p <.001$. Overall, participants reported that they enjoyed playing Mario Kart the most (56% of responses) and the average player engaged in very little game play, with 39% stating that they never play video games.

3.3.3 Discussion

The majority of MW research has been conducted using traditional laboratory tasks (e.g., simple attention paradigms), thus the purpose of this experiment was to investigate MW as it applies to a real-world, engaging context -- playing video games -- one in which low levels of MW would be expected. I was also interested in how different visual platforms could be used as a way to manipulate attentional engagement (MW rates, interest, and engagement) within video game play. Contrary to the hypothesis, none of MW rates or any of the other subjective ratings were influenced by screen size (projector screen, computer monitor, and handheld device). Furthermore, based on Bayes analyses, these screen conditions are more similar than dissimilar. Even though my primary hypothesis was not supported, the present experiment provides evidence that MW is a regular occurring cognitive state that persists, occurring on average 20% of the time within engaging, real-world activities, independent of screen size.

This experiment holds practical implications when considering tasks that require vigilant attention to screens, and have negative costs associated with MW, such as vehicle operations or security industries (He et al., 2011). By determining the circumstances under which screen size influences MW rates could inform safety and strategy protocol in tactic to reduce MW rates. The findings of the present experiment may also have implications for consumer-oriented industries, such as educational institutions, marketing and advertising firms. Furthermore, evidence suggests that screen content may in fact be more successful at capturing viewers' attention than increasing screen size (Ivory & Magee, 2009; D. Kim & Kim, 2012; Reeves et al., 1999; Unsworth & McMillan, 2013). Thus, the results of prior work that "bigger is better" (for attention) during passive watching may not necessarily extend to rates of attention during more active tasks, such as video game play, which requires greater physical and cognitive demands than passively viewing. Burns and Fairclough (2015), for example, report that screen size does not impact physiological or subjective immersion of game play. In another experiment, character impression, mood, and physical presence were more positive for a larger screen than a smaller one, but screen size did not influence self-presence or immersion (Hou et al., 2012). Together with the present experiment, this suggests that strategic changes to improve screen content may increase engagement and decrease MW rates. Such an outcome would be beneficial when targeting products – as a wider audience could be reached if a better experience did not have to mean a bigger (and likely more expensive) product.

An unexpected but meaningful outcome of the present experiment was the finding that the game played had a significant impact on some of the dependent variables. The games were selected based on their current popular status, with the expectation that they would all be classified as equally interesting and engaging. The outcome of task features impacting results, above and beyond those being manipulated, demonstrates another critical factor that researchers should consider when conducting MW research. Researchers typically use the same materials across task manipulations (i.e., one passage or lecture video) without considering whether the results would generalize when different materials are used. Thus, although unexpected, the present experiment presents an important case for manipulating task characteristics as well as measuring interest in content prior to testing a planned manipulation.

This experiment is not without limitations. The fact that 39% of participants never play video games indicates that gaming is not an all-inclusive leisure activity. That said, by selecting highly popular games, I hoped that even those who were not active gamers would enjoy the task more so than a traditional laboratory sustained attention task. By limiting the study of MW in leisure to video games, I obtained more experimental control than would be achieved by allowing each participant to choose their own leisure activity. Future research could examine personal preference by offering a limited selection of leisure activities for the participant to choose from, allowing for a higher likelihood of regular engagement with the task. Through investigating different types of leisure activities, I would gain a more thorough representation of the range of MW rates that people experience in leisure activities.

Despite a wealth of research on MW rates within lab-based settings, MW rates during naturalistic tasks are almost exclusively limited to reading or lecture-viewing tasks. The present experiment engaged participants in a more arousing and less repetitive task (as compared to standard laboratory tasks), yet found that what could be thought of as the lower-bound of potential MW rates, was still well above zero. Specifically, the average rate of MW during video game play was 22%, a rate that deviates from the often cited 30-50% rate, but is still not singularly exceptional. Comparable rates, though less common, have been reported in the past with some in-lab experiments using reading or basic sustained attention tasks (26%: Franklin et al., 2014, 2013; 21%: Jackson & Balota, 2012; 19%: Sayette, Schooler, & Reichle, 2010). Thus, it appears that reducing mind wandering below 20% might be a difficult feat for any lab-based experiment and the standard trend of mind wandering increasing overtime holds even in tasks that are known to be highly engaging in the real-world.

3.4 Chapter summary

The present chapter included three experimental examinations of MW within naturalistic tasks or settings. Experiments 6 and 7 examined the question of how manipulated features of the classroom setting (i.e., instructor visibility and peer presence) impact student learning and MW rates. Experiment 8 investigated what the lower bound of MW might be, given the use of a naturalistic and enjoyable leisure task, playing a video game.

Experiment 6 tested how manipulation of physical instructor presence, while holding all other factors of a live learning environment constant, would impact MW rates, learning and interest and motivation ratings. The results indicate that having a live (physically present) instructor leads to better lecture retention, and greater interest and motivation ratings. This investigation opens the door for more research in applied settings, using controlled experimental procedures, to investigate other factors that differ between live and online settings. For example, can we augment the virtual setting to overcome any negative impact that might weaken it compared to a live session?

Experiment 7 built on the findings of the previous experiment by examining whether other factors that differs between settings (instructor visibility and peer presence) impact cognitive performance. Although online lectures are an increasingly popular tool for learning, research on instructor visibility during an online lecture, and students' environmental settings, has not been well-explored. Through experimentally manipulating online display format and social learning settings, I found evidence that students' social setting during viewing has an impact on MW, with students who watched the lecture in a classroom with others reporting significantly more MW than students who watched the lecture alone. Social setting also moderated the negative relationship between MW and material retention. Overall, these results demonstrate that learning experiences during online lectures can vary based on where, and with whom, the lectures are watched.

The final experiment in this chapter investigated whether different visual platforms (a handheld device, a computer monitor, and a projector screen) would influence MW reports during a highly natural task, video-game play. Results suggest that gaming platform does not influence MW rates, mental and physical demand, interest, or attention ratings. That said, the observed MW rates of 20% converge with the lower bound occasionally reported in the field.

Chapter 4: Mind Wandering and Distraction

When reading a novel on the bus one's attention may shift away from reading to thinking about things such as what to prepare for dinner, or noticing the ring tone of a nearby passenger's phone. Such shifts of attention can be either inward, toward one's own thoughts (i.e., MW), or outward, toward an external event (i.e., a distraction). In our daily lives the environments in which we process information can vary greatly (e.g., the number of people, peripheral sights and sounds, and secondary tasks) and these variations may influence lapses of attention in different ways. The final chapter of my thesis directly pits different task settings (i.e., research laboratory versus everyday life) against each other to measure their potential influence on MW, distraction rates, and retention performance.

As discussed in the introduction, higher rates of MW are often associated with greater detriments to performance (e.g., slower reaction times, greater errors, poorer memory retention; Dixon & Li, 2013; Forster & Lavie, 2009; McVay & Kane, 2010; Smallwood, McSpadden, & Schooler, 2008; Varao Sousa et al., 2013). But task performance can also be negatively impacted by another form of inattention: distraction (e.g., poorer memory retention, slower reaction times; Banbury & Berry, 1998; Forster & Lavie, 2007; Reinten, Braat-Eggen, Hornikx, Kort, & Kohlrausch, 2017). The impact of, and relation between, MW and distraction within a single task suggests increased rates of MW and distraction during the SART are related to increased response variability and errors (Stawarczyk, Majerus, Maj, Van der Linden, and D'Argembeau, 2011). Additional research by Unsworth and colleagues (Robison & Unsworth, 2015; Unsworth, Brewer, et al., 2012; Unsworth & McMillan, 2014; Unsworth, McMillan, et al., 2012) suggests that MW and distraction are separate constructs, with distractions interfering with performance "over and above that accounted for by general lapses of attention" (Unsworth & McMillan, 2014,

p. 23). Collectively, this research suggests that both distractions and MW influence performance (but see Olivers & Nieuwenhuis, 2005).

What remains relatively unclear, however, is whether the effects found in the laboratory settings described above fully capture the role of MW and distractions in everyday life. While some researchers have tried to address concerns about unnatural laboratory settings by using less controlled, more naturalistic distraction stimuli (i.e., a background television program, or restaurant audio track; Pool, Koolstra, & Voort, 2003; Robison & Unsworth, 2015), these distractions are still limited to scheduled experimental events from a limited stimulus set. In day-to-day life, distractions are often more unpredictable than those experienced in a laboratory experiment.

Notably, some experiments have revealed important differences in patterns of inattentiveness across laboratory and everyday settings (Kane et al. 2017; Wammes & Smilek, 2017). For instance, Kane et al. (2017) measured MW rates in the laboratory while participants completed a variety of executive-control paradigms (e.g., the sustained attention to response task, the attention flanker task). These results were compared to results obtained in everyday life using experience sampling methods whereby participants responded to thought-probes on a handheld device. Two key points of divergence between MW in the laboratory and in life were observed. First, MW rates during laboratory tasks did not correlate with MW rates collected during everyday tasks, leading the authors to conclude that "any relation between laboratory and overall daily-life mind-wandering propensities is not robust." (p. 1278). Second, while measures of executive control were related to MW rates in the laboratory, they were not directly related to MW rates outside of the laboratory.

The divergence between life and lab is also supported by a recent experiment by Unsworth and McMillan (2017) in which participants completed in-lab cognitive ability tests and then, for one week, reported MW and distractions that took place while studying or while in class. In contrast to laboratory experiments showing that MW and distraction negatively influence performance (Banbury & Berry, 1998; Dixon & Li, 2013; Mcvay & Kane, 2010; Smallwood et al., 2008; Varao Sousa et al., 2013), their results suggest that everyday reports of MW and distraction were not correlated with in-lab cognitive ability measures, or with academic performance. It is worth noting that there were exceptions in Unsworth and McMillan (2017), such that some categories of MW and distraction did correlate with in-lab measures (e.g., "mind wandering due to disinterest" correlated with most in-lab tasks, while "Distraction due to hunger" did not) (see also Kane et al. 2017; Kane et al. 2007; Unsworth et al., 2012). These results suggest that the negative consequence of attentional failures is not absolute, and varies with task setting. This dovetails with the idea that the true diversity and impact of inattention in everyday environments may not have been captured by previous experiments of MW and distraction in the lab.

Studies of MW during university lectures also show different patterns of MW between lab and life settings. When students are asked to watch video recordings of lectures in the laboratory, MW rates typically increase as the lecture progresses (Farley et al., 2013; Risko, Anderson, et al., 2012; Wammes & Smilek, 2017). In contrast, when attending a live lecture, MW rates of undergraduates remain stable across the lecture (Wammes, Boucher, et al., 2016; Wammes & Smilek, 2017). These findings suggest that patterns of inattention may change across lab and natural situations.

Given recent results indicating a divergence in MW rates across laboratory and real-life situations, determining why such a divergence occurs becomes an important issue. One plausible reason for the divergence, which the next two experiments focus on, concerns the availability of distractions in the laboratory and in everyday life. In the laboratory, testing situations are typically controlled, with very few distractions available to co-opt attentional resources. In fact, laboratory settings are specifically designed to reduce and control distraction. In such cases, MW is often the primary way (if not the only way) that inattention can be manifested. In contrast, within the natural world, opportunities for distraction seem frequent and diverse (e.g., the presence of other people, peripheral sights and sounds, media devices). In such a complex setting inattention may be expressed in ways other than MW. If the rate and impact of inattention differ between the lab and the outside world, previous lab research may be underestimating the impact that inattention has in uncontrolled environments.

To determine whether rates of inattention differ when inside and outside of the lab, I conducted two experiments that manipulated whether participants listened to an audiobook while inside the lab (controlled setting) or outside the lab (uncontrolled setting). Participants were allowed to move freely outside the lab to reduce the constraints of a lab testing room, and amplify the “noise” due to the perceptual richness or affordances found in everyday environments. An audiobook task was chosen for two reasons: 1) mind wandering occurs frequently and reliably in this setting (Kopp & D’Mello, 2015; Varao Sousa et al., 2013), and 2) an audiobook allows one to move freely while still completing the primary task of listening. The decision to allow participants to move freely while outside the lab was in keeping with the fundamental methodology of cognitive ethology (Kingstone et al., 2005, 2008; Smilek et al., 2006). The goal was to first discover what people naturally do in an everyday environment and how that impacts cognition. Thus participants were not instructed on how they should behave outside the lab, anticipating that choices made would reflect normative behavioural responses.

The goal of both experiments in this final chapter of my thesis was to examine how inattention rates (MW and distraction) are affected by different task environments, specifically, inside a quiet lab versus outside in a dynamic real-world setting. In addition to examining how different forms of inattention are influenced by task setting, I also examined whether inattention rates were influenced by reporting method (self-caught versus probe-caught). Experiment 9 uses the standard probe-caught paradigm with participants prompted by an audio tone, while listening, to report whether they were on task, mind wandering, or distracted. Experiment 10 investigated whether rates of MW and distraction would differ when measuring the rates via self-caught methodology. While the experiments in Chapter 2 of my thesis do not suggest fundamental differences in rates of MW between the two methods, it seemed prudent to investigate the pattern of self-caught rates for different *types* of inattention.

It was predicted that rates of inattention would not be equivalent between life and lab environments. Specifically, I predicted participants outside the lab would report more instances of distraction, relative to MW, than those inside the lab, as the former affords exposure to a greater array of dynamic stimuli. Since MW is an internally driven process there was no reason to predict that different external environments would influence the rates of MW. Since Unsworth and McMillan (2014) found that distractions impaired retention test performance, and I predicted greater distraction outside the lab, I further predicted that retention test performance should be worse when the audiobook was listened to outside the lab than inside. I also conducted an exploratory analysis to see if self-reported ratings of interest, motivation or boredom (factors commonly correlated with MW) were influenced by the task setting.

4.1 Experiment 9: Probe-caught MW and distractions

4.1.1 Methods

4.1.1.1 Participants

Sample size (*n* = 34/group) was determined by conducting a power analysis using G*Power software (Faul et al., 2007). This analysis was based on the effect size (*d* = .78) found in Robison and Unsworth (2015), who looked at differences in MW and distraction in a laboratory setting. Seventy-three students, 12 men and 61 women, aged 18 to 42 (*M* = 21.32, *SD* = 3.57) from the University of British Columbia participated in the experiment. Participants were compensated with course credit. Eight participants were dropped from the analysis due missing responses (i.e., did not respond to all the probes), leaving 34 in the Inside (controlled) setting and 31 in the Outside (uncontrolled) setting.

4.1.1.2 Audiobook

A 14 minute excerpt from *Weapons of Math Destruction* (O'Neil, 2016) narrated by the author was used. The audiobook was presented on a 4th generation iPod Shuffle, with a generic pair of headphones.

4.1.1.3 MW probes

Ten audio tones were dispersed roughly 1-2 minutes apart throughout the audio track. A sample of the audio probe was presented to participants during the instructions. Participants were provided a numbered response sheet and responded to the audio probes by circling one of the following options after each tone: "I am focused on the audiobook", "I am distracted by sights/sounds", and "I am mind wandering".

4.1.1.4 Retention test

A retention test on the contents of the audiobook was administered via an online Qualtrics survey. The retention test consisted of nine multiple choice questions and one short answer question.

4.1.1.5 Other measures

Ratings of motivation to attend the audiobook and interest in the material were collected using a five-point Likert scale. Boredom, as a state of disengagement, was measured via the seven-point Likert "Short Boredom Proneness Scale" (Struk et al., 2017). Participants were also asked about their experience during the task (what they did while listening; whether the task reflected real-world audiobook listening; and how often they listened to audiobooks). Participants assigned to the Outside condition were additionally asked where they went, what proportion of time they spent outside, and if they interacted with anyone. All questions were administered via Qualtrics.

4.1.1.6 Procedure

After providing informed consent, participants were randomly assigned to either the Inside or Outside condition. Participants assigned to the Inside condition remained inside a laboratory testing room for the duration of the audiobook track, while participants assigned to Outside condition were allowed to move freely anywhere outside the laboratory. The ipod "lock" feature was enabled so that participants could not pause, rewind or skip through the audio track. Participants in the Outside condition were advised to return to the lab once they had heard and responded to the tenth and final MW probe, while continuing to listen to the audiobook. Participants in the Outside condition who arrived with time remaining on the audiobook track were taken to the testing room to listen to the remainder of the track. Once the audiobook track had finished, the Qualtrics survey was administered.

4.1.2 Results

All statistical analyses are reported with two-tailed *p*-values. In cases where the samples were not homogeneous, as indicated by Levene's test of homogeneity, an adjusted degrees of freedom value is reported. Overall inattention was calculated by summing MW and distraction rates. Table 4.1 shows descriptive statistics for all variables measured.

Measure	Inside (N = 34)	Outside (N = 31)
Distraction rate – proportion (Range 0-.5)	.18 (.17)	.24 (.12)
MW rate – proportion (Range 0-.6)	.20 (.13)	.14 (.10)
Overall inattention – proportion (Range 0-.7)	.38 (.19)	.38 (.18)
Retention test score – proportion (Range 0.18-1)	.62 (.16)	.50 (.19)
Interest (Range 1-4)	2.35 (.77)	2.26 (.82)
Motivation (Range 1-5)	3.03 (1.09)	3.26 (.89)
Boredom (Range 1.6-5.6)	3.26 (.96)	3.01 (.90)

Table 4.1. Mean reports of dependent measures, with standard deviations in parentheses.

4.1.2.1 Inattention ratings

A 2x2 between-within ANOVA assessed whether the condition (Inside or Outside) influenced MW or distraction ratings[7]. As indicated in Figure 4.1, the between-subjects test indicated that overall inattention did not differ when Inside versus Outside, $F(1, 63) = .003$, $p = .96$, $BF_{01} = 5.32$, $\eta p^2 < .001$. There was also no within-subjects main effect of report type, $F(1, 63) = 2.30$, $p = .14$, $BF_{01} = 1.59$, $\eta p^2 = .04$. However, an interaction revealed that differences between specific inattention rates (i.e., distraction or MW) varied with condition, $F(1, 63) = 6.75$, $p = .012$, $BF_{10} = 1.15$, $\eta p^2 = .10$. Figure 4.2 illustrates the follow up, paired t-tests which indicate that MW and

[7] I recognize that the reports are partially ipsative in nature given the three forced-choice alternative (mind wandering, distracted, on-task). I nonetheless made an apriori decision to focus our analysis on differences between types of inattention and thus follow the protocol of similar analyses (Robison & Unsworth, 2015; Unsworth & McMillan, 2014)

distraction rates did not differ for those Inside, $t(33) = .57$, $p = .57$, 95% CIs [-11, .06], $BF_{01} = 3.38$, $d = .099$, however participants Outside reported significantly more inattention due to distraction than MW, $t(30) = 4.41$, $p < .001$, 95% CIs [.06, .15], $BF_{10} = 58.26$, $d = .81$.

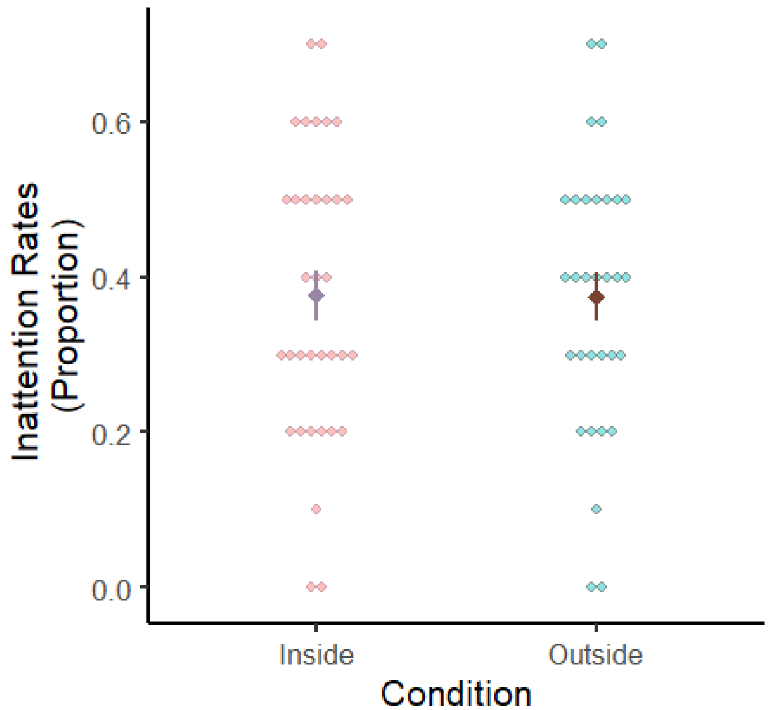


Figure 4.1. Inattention reports across conditions. Bars represent one standard error of the mean.

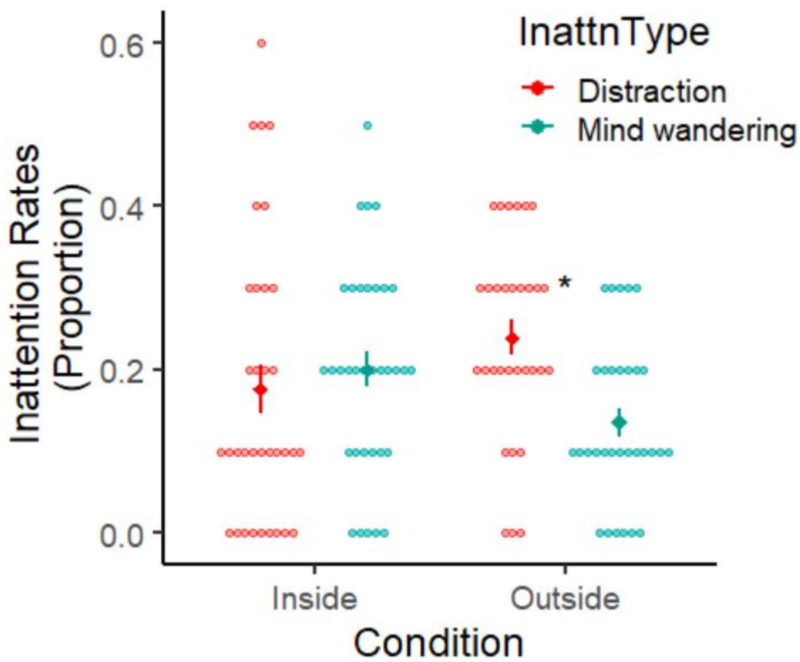


Figure 4.2. Breakdown of types of inattention reports across conditions. Bars represent one standard error of the mean.

4.1.2.2 Retention test performance

There was a significant difference in retention based on condition, such that those Inside performed significantly better than those Outside, $t(59.33) = 2.78$, $p = .007$, 95% CIs [.03, .20], $BF_{10} = 6.02$, $d = .69$.

4.1.2.3 Interest, motivation and boredom scales

None of the Interest ($t(62) = .48$, $p = .63$, 95% CIs [-.30, .48], $BF_{01} = 3.57$, $d = .12$), Motivation ($t(62) = .93$, $p = .36$, 95% CIs [-.72, .26], $BF_{01} = 2.77$, $d = .23$) or Boredom ($t(62) = 1.09$, $p = .28$, 95% CIs [-.21, .72], $BF_{01} = 2.38$, $d = .27$) ratings differed between the Inside and Outside conditions.

4.1.2.4 Audiobook listening frequency

When asked how often they listen to audiobooks, 24 participants stated “Never”; 28 reported very seldom or rarely listening; 7 stated a couple times per year and 6 said once per month or of a greater frequency. For those who reported listening to audiobooks more than “Never”, I asked if the task reflected how they would normally listen to an audiobook. Sixty-one percent of those in the Inside condition (n = 23) and fifty-six percent of those in the Outside condition (n = 18) reported that it did. For those who felt that the task did not reflect their normal audiobook listening experience, the most common reason given was that they normally did something additional (chores, driving, on a bus, etc.) while listening.

4.1.2.5 Other variables

Both groups were asked what they did while listening to the audiobook. A majority of those in the Outside condition indicated walking around inside the building or sitting down, while those in the Inside condition mainly reported looking around the room. Outside participants also reported how long they spent outside, however responses were not clearly identified as outside the laboratory or outside the building, and so this question could not be assessed further. A majority (n = 25) of those in the Outside condition stated that they did not interact with anyone, a few (n = 5) reported brief interactions, and one individual reported having a five-minute conversation.

4.1.3 Discussion

The present experiment examined how the rates of two forms of inattention – MW and distraction – vary with environment, i.e., an uncontrolled natural setting and a controlled laboratory setting. These results suggest that overall rates of inattention are similar when one is outside in a dynamic real-world setting or inside a controlled laboratory, but there are differences in whether inattention manifests as MW or as distraction. While there were no differences between MW and distraction when inside the lab (convergent with Robison & Unsworth 2015), participants outside reported more distraction than MW (convergent with daily diary studies: Unsworth, Brewer, et al., 2012; Unsworth & McMillan, 2017). A secondary goal was to examine whether inattention rates as a function of setting differentially impact retention performance. As predicted, distraction rates were higher and retention test performance was worse for those outside than inside the laboratory.

Motivated by recent research suggesting that results found in the laboratory may not generalize to everyday behaviour (Kane et al. 2017; Unsworth & Mcmillan, 2017; Wammes & Smilek, 2017), the present experiment examined how MW and external distractions differ between a dynamic, out-of-lab setting and a controlled, in-lab setting. In line with Unsworth and colleagues (Robison & Unsworth, 2015; Unsworth & McMillan, 2014; Unsworth, McMillan, Brewer, & Spillers, 2012), the current experiment substantiates the finding that MW and external distractions are two unique types of attentional failures. Furthermore, the allocation of inattention depends on the environment: Distraction was experienced more often than MW outside the laboratory, while the rates of distraction and MW were equivalent inside the laboratory. Additionally, retention performance was worse when outside than inside the lab. The next experiment examines whether the results found in the present experiment (discrepant inattention and retention rates between a laboratory environment and a dynamic real-world environment) generalize to the self-caught paradigm.

4.2 Experiment 10: Self-caught MW and distractions

The experiments in Chapter 2 of my thesis suggest that MW rates are not impacted by whether they are reported via probe or self-caught reports. In the present experiment, I reconsider the potential for inattention rates to be impacted by report type. Specifically building on Experiment 9, this experiment investigates whether self-caught reports of inattention (MW and distraction) differ between settings (controlled lab versus uncontrolled out-of-lab setting). If, as supported by the experiments in Chapter 2, self-caught methods capture inattention in much the same way as probe-caught methods, then I would expect that the pattern of results would mimic Experiment 9. That is to say, individuals would self-catch more distraction outside the lab than MW, and inside the lab the frequency of these reports would not differ. In contrast, if the self-caught method captures a different type of attentional lapse, then the pattern of self-caught distractions may differ. In addition to investigating the pattern of results found with the self-caught method, I also conduct an exploratory analysis that directly compares the results to those already obtained in Experiment 9, to provide further evidence for differences between self-caught and probe-caught methodologies.

4.2.1 Method

4.2.1.1 Participants

Ninety-three students, 19 men and 74 women, aged 18 to 44 (M = 20.92, SD = 3.99) from the University of British Columbia participated in the experiment. Sample size (n = 34/group) was determined based on the effect size (d = .78) reported for external distraction report differences between the "noise" and "silence" condition in Robison and Unsworth (2015), with an extra two weeks of data collected to ensure that sufficient sample was obtained after exclusions. Participants were compensated with course credit. Seven participants were dropped from the analysis due missing responses (i.e., lost wi-fi connection while completing Qualtrics survey), leaving 46 in the Inside (controlled) setting and 40 in the Outside (uncontrolled) setting.

4.2.1.2 Audiobook

A 14 minute excerpt from *Weapons of Math Destruction* (O'Neil, 2016) narrated by the author was used, same as Experiment 9 (Chapter 4). The audiobook was presented on a smartphone, with a generic pair of headphones.

4.2.1.3 MW reports

Participants were provided with an online Qualtrics survey via a smartphone to record their responses. Participants were asked to indicate anytime they noticed that they were MW or distracted by pressing the corresponding button on the phone survey. The response buttons included "I am distracted by sights/sounds", and "I am mind wandering". After listening to the first 10 minutes of the audiotrack, an auditory prompt sounded, which indicated that participants should close the survey but continue listening to the audiobook. Participants in the outside condition used this cue to signal that they should return to the lab. Participants in the inside condition were advised to continue listening until the researcher returned to the room (after 14 minutes).

4.2.1.4 Retention test

A retention test on the contents of the audiobook was administered via an online Qualtrics survey. The retention test consisted of the same questions as in Experiment 9.

4.2.1.5 Other measures

A five-point Likert scale was used to collect ratings of motivation to attend the audiobook and interest in the material. Boredom was measured via the seven-point Likert "Short Boredom

Proneness Scale" (Struk et al., 2017). Participants were also asked about their experience during the task (behaviour while listening; real-world nature of the task; audiobook listening frequency). Participants assigned to the Outside condition were additionally asked where they went, and if they had an social interaction.

4.2.1.6 Procedure

After providing informed consent, participants were randomly assigned to either the Inside or Outside condition. For the Inside condition, participants remained inside a laboratory testing room for the task, while participants assigned to Outside condition were allowed to move freely anywhere outside the laboratory while listening to the audiobook. Participants were asked to not pause, rewind or skip through the audio track. Participants in the Outside condition were advised to return to the lab once they had heard an in-track audio prompt, while continuing to listen to the audiobook. Once the audiobook track had finished, the Qualtrics survey was administered inside the lab testing room.

4.2.2 Results

All statistical analyses are reported with two-tailed *p*-values. In cases where the samples were not homogeneous, as indicated by Levene's test of homogeneity, an adjusted degrees of freedom value is reported. Overall inattention was calculated by summing MW and distraction rates. Table 4.2 shows descriptive statistics for all variables measured.

Measure	Inside (n = 46)	Outside (n = 40)
Distraction rate (Range 0-14)	2.17 (2.39)	5.05 (3.58)
MW rate (Range 0-17)	4.02 (3.26)	5.23 (3.45)
Overall inattention (Range 0-25)	6.20 (5.00)	10.28 (6.17)

Retention test performance (%) (Range 0-100)	.66 (.24)	.65 (.25)
Interest (Range 1-4)	2.59 (.93)	2.05 (.95)
Motivation (Range 1-5)	3.13 (1.20)	3 (1.26)
Boredom (Range 1.13-6.25)	3.27 (1.12)	3.54 (1.13)

Table 4.2. Mean reports of dependent measures, with standard deviations in parentheses.

4.2.2.1 Inattention ratings

A 2x2 between-within ANOVA assessed whether the condition (Inside or Outside) influenced MW or distraction ratings[8]. The between subjects test indicated that overall inattention differed when Inside versus Outside, $F(1, 83) = 10.94$, $p = .001$, $BF_{10} > 150$, $\eta p^2 = .12$ (Figure 4.4). There was also a within-subjects main effect of report type, $F(1, 83) = 10.38$, $p = .002$, $BF_{10} = 12.71$, $\eta p^2 = .11$. This was further qualified by an interaction revealing that differences between specific inattention rates (i.e., distraction or MW) varied with condition, $F(1, 83) = 5.94$, $p = .017$, $BF_{10} > 150$, $\eta p^2 = .07$. Follow up, paired t-tests indicate that MW rates were significantly higher than distraction rates for those Inside, $t(45) = 4.51$, $p < .001$, 95% CIs [-2.67, -1.02], $BF_{10} = 13.33$, $d = .66$, however for Ps outside, MW and distraction rates did not differ significantly, $t(39) = .33$, $p = .75$, 95% CIs [-1.25, .90], $BF_{01} = 4.21$, $d = .05$.

[8] We recognize that the reports are partially ipsative in nature given the three forced-choice alternative (mind wandering, distracted, on-task). We nonetheless made an apriori decision to focus our analysis on differences between types of inattention and thus follow the protocol of similar analyses (Robison & Unsworth, 2015; Unsworth & McMillan, 2014)

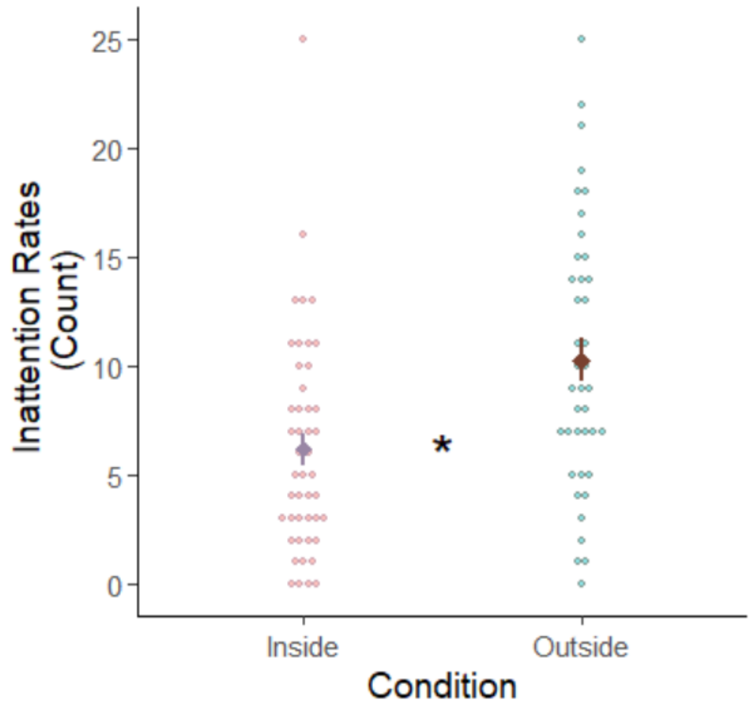


Figure 4.3. Overall inattention reports across conditions. Bars represent one standard error of the mean.

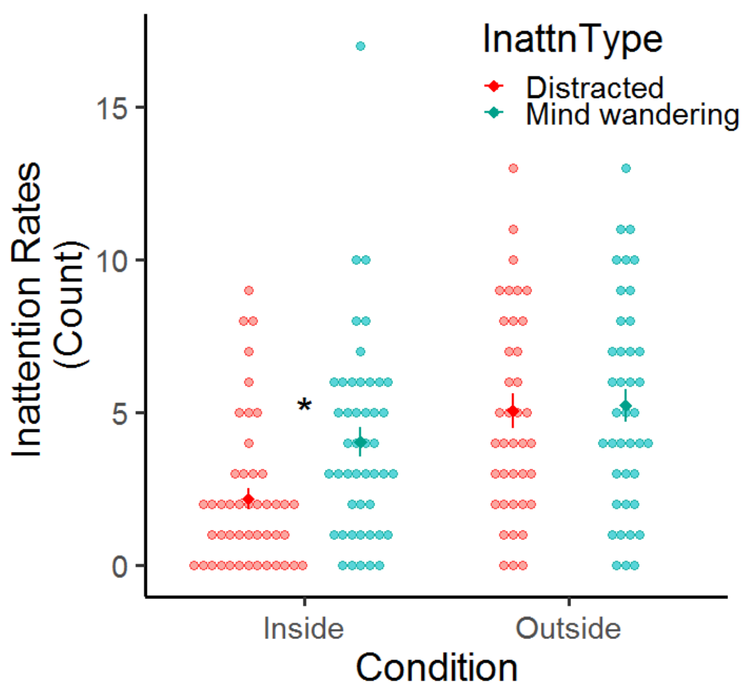


Figure 4.4. Inattention rates split by condition and attention type. Bars represent one standard error of the mean.

4.2.2.2 Retention test performance

There was no significant difference in retention test performance based on condition, $t(82) = .23$, $p = .82$, 95% CIs [-.83, 1.05], $BF_{01} = 4.35$, $d = .05$.

4.2.2.3 Interest, motivation and boredom scales

Interest was rated as significantly higher for those inside the lab than those outside, $t(82) = 2.62$, $p = .01$, 95% CIs [.13, .94], $BF_{10} = 4.34$, $d = .49$. Neither Motivation, $t(82) = .49$, $p = .63$, 95% CIs [-40, .66], $BF_{01} = 4$, $d = .11$, nor Boredom ratings, $t(82) = 1.12$, $p = .27$, 95% CIs [-.75, .21], $BF_{01} = 2.56$, $d = .24$, differed between the Inside and Outside conditions.

4.2.2.4 Audiobook listening frequency

When asked how often they listen to audiobooks, 23 participants stated “Never”; 35 reported very seldom or rarely listening; 14 stated a couple times per year and 21 said once per month or of a greater frequency. For those who reported listening to audiobooks more than “Never”, I asked if the task reflected how they would normally listen to an audiobook. Only 28% of those in the Inside condition ($n = 11$) and 35% of those in the Outside condition ($n = 14$) reported that it did. For those who felt that the task did not reflect their normal audiobook listening experience, the most common reason given was that they normally did something additional (chores, driving, on a bus, etc.) while listening or would be in a more comfortable position.

4.2.2.5 Other variables

The majority of those in the Outside condition indicated walking around inside the building and looking around or sitting down inside the building, while those in the Inside condition mainly reported looking around the room or listening with their eyes closed. Nearly all of those in the Outside condition stated that they did not interact with anyone, with only 2 reporting brief interactions.

4.2.2.6 Additional analyses

A secondary goal of this experiment was to examine whether self-caught and probe-caught inattention reports differ between the Inside and Outside setting. To conduct this analysis, the inattention data from Experiment 9 was compared to that of the present experiment with Inattention Reports (MW, Distraction) calculated as a function of overall inattention – much like the analyses in Experiments 2 and 3 in Chapter 2 – thus putting the self-caught and probe-caught rates on equal footing. Figure 4.5 visually represents this comparison. As can be seen, when participants self-catch, the differences between MW & distraction manifests in the Inside group, however when participants are reporting via probe-catching, the differences between MW & distraction manifest in the Outside group. It is important to note that this analysis was not planned in advance of collecting data for each of the experiments, which were conducted one year apart. Thus, it would be prudent for further experiments to examine whether cohort differences might exist and additional within-subjects designs could explore the stability of differences between self-caught and probe-caught reporting across settings varying in naturalness.

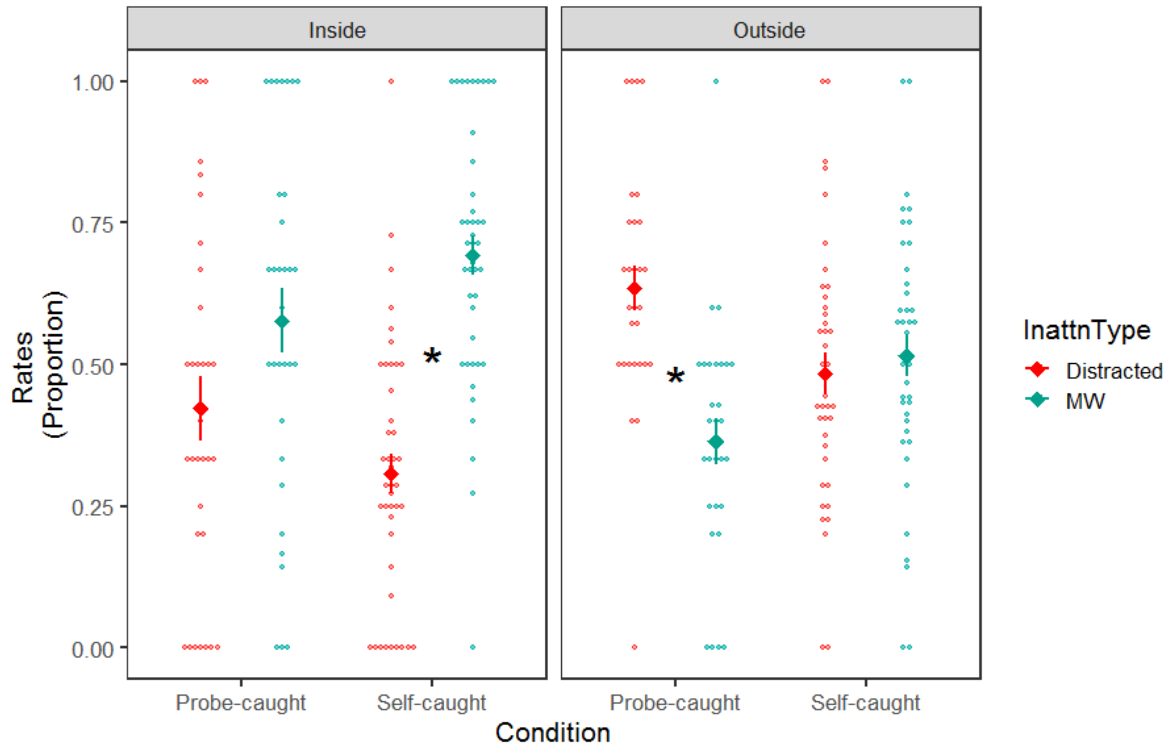


Figure 4.5. Inattention rates split by experiment, condition and attention type. Bars represent one standard error of the mean.

4.2.3 Discussion

The present experiment extended the examination of differences between two types of inattention (MW and distraction) within natural and unnatural environments to a self-caught methodology. Contrary to Experiment 9 (Chapter 4), where a probe-caught methodology was used, there were significant differences in overall rates of inattention, with participants reporting more inattention outside in a dynamic real-world setting than inside a controlled laboratory. Furthermore, there were differences in whether inattention manifests as MW or as distraction, with participants inside reporting more MW than distraction, and no differences between MW and distraction when outside the lab (contradictory to Robison & Unsworth 2015; Unsworth, Brewer, et al., 2012; Unsworth & McMillan, 2017). Also contradictory to the results in

Experiment 9, when using the self-caught method differences in retention were not found between the two conditions, despite there being significant differences in overall inattention between conditions.

The present experiment applies the fundamental methodology of cognitive ethology (Kingstone et al., 2005, 2008; Smilek et al., 2006) by exploring the nature of inattention in a natural everyday environment and by using a more naturalistic reporting method, the self-caught report, to capture these lapses of inattention. In summary, rates of inattention do differ when measured inside versus outside the lab, thus motivating continued work comparing generalizations of laboratory to everyday behaviour (Kane et al. 2017; Unsworth & Mcmillan, 2017; Wammes & Smilek, 2017). Furthermore, it seems that allocation of inattention depends on both the environment and the report method, such that distraction was experienced more often than MW outside the laboratory when probe-catching, but MW was experienced more often than distraction inside the laboratory when self-catching.

4.3 Chapter summary

Attention research has found that performance between lab and everyday settings are not always comparable (Kingstone et al., 2008, 2003). For example, looking behavior differs when in social real-world settings as compared to laboratory settings (Foulsham, Walker, & Kingstone, 2011; Laidlaw et al., 2011). In this chapter, my goal was to examine how MW and distraction rates differ between lab and real-world settings. The inclusion of distraction as another measure of inattention was motivated by recent attention research, which has examined how varying types of inattention occur in everyday settings. Specifically, through a daily-diary study Unsworth, McMillan, Brewer, and Spillers (2012) found that distraction and MW while studying were the most highly reported forms of inattention during everyday tasks for an undergraduate student population.

The final chapter of my thesis included two experiments designed to extend findings of Robison and Unsworth (2015) by further pushing the limit on what constitutes a “natural” setting. The two experiments that follow attempted to move beyond a simple lab-based manipulation of “noise” and insert participants into a more dynamic and natural environment with participants leaving the lab and having freedom to explore the nearby environment. The goal of both experiments was to examine how inattention rates (MW and distraction) are affected by different task environments, as well as how different forms of inattention are influenced by reporting method. Results from both experiments suggest that research taking place within a laboratory environment may not reflect the true impact of dynamic real-world environments on inattention. This outcome is supported by prior research showing discrepancies in MW reports between lab and daily life (Kane et al. 2017; Unsworth & McMillan, 2017;Wammes & Smilek, 2017), however is the first to illustrate that the lab-life discrepency may relate to both reporting type (self-caught versus probe-caught) as well as inattention type (MW versus distraction). Future work would benefit from a re-examination of MW differences within lab and daily life to see what role distraction has in contributing to prior divergent rates between settings.

Future work would benefit from moving beyond MW and external distraction in the context of listening to audiobooks, to examine inattention in the context of other everyday tasks, such as completing homework, driving, or preparing dinner, settings that also afford unconstrained attention. Although prior research has examined inattention in the context of complex tasks such as driving and real-world learning settings (e.g., Cowley, 2013; Dingus et al., 2016; Unsworth & McMillan, 2017; Yanko & Spalek, 2013), these experiments have not examined the distinct roles of MW and distraction. With greater task freedom, individuals may be able to anticipate or choose how to navigate their surroundings and thus optimize their task experience, thereby modulating the consequences of inattention.

One limitation to the present experiments is that it is unclear to what extent the behavior observed outside the laboratory truly represents real life behaviour. Many audiobooks users reported that the task did not completely reflect their “real-life” audiobook listening, noting that a concurrent task would normally be completed while listening (e.g., driving, chores, etc.). This suggests that the present uncontrolled, out-of-lab setting may still not be entirely capturing real life behaviour, as it lacked the natural conditions under which participants would normally complete the audiobook listening task. It would be prudent to conduct follow-up experiments that include a series of settings that reflect common audiobook listening contexts. Results from such a design would reveal whether, when within a more natural setting, inattention rates deviate from those found here, and why (e.g., the role of personal control, task familiarity, motivation, etc.).

Chapter 5: General Discussion

Over the past two decades researchers have actively pursued an understanding of the nature of a wandering mind. Yet, the majority of this research has taken place within tasks and settings that are unnatural (i.e., simple lab-based paradigms) and do not reflect how individuals behave in everyday life. The goal of this thesis was to examine MW rates in a variety of natural tasks and settings, and answer the following research questions: (i) whether the way in which one reports MW (self-caught versus probe-caught) influences rates or report content; (ii) how MW rates vary when presentation style or task are manipulated; and (iii) whether MW and distraction occur at different rates within lab and non-lab settings. In this concluding chapter I will first summarize the implications of the results found, and then discuss the limitations of the present findings and interpretations.

5.1 Thesis summary

Across ten experiments, I examined the rate, characteristics, and impact of MW in tasks designed to imitate everyday activities (e.g., reading a novel, playing a video game) or settings that reflect natural behaviour (classrooms, non-lab environments). The experiments included in Chapters 2-4, were designed help address specific research questions. Beginning with Chapter 2, the primary goal of the first five experiments was to determine whether the way in which one reports MW (self-caught versus probe-caught) influences rates or report content. Experiment 1 provided evidence that the format of a MW probe (audio or visual) did not impact MW rates or performance, even when the probe format was incongruent with the task itself (i.e., a visual probe with an auditory task). The primary aim of experiments 2-4 was to examine characteristics of the wandering mind beyond a simple 'on-' or 'off-task' response. Results from Experiment 2 suggest that intentional MW is more likely to occur while watching a lecture than while reading a passage (a passive versus active task), but that overall unintentional MW is far more prominent than intentional MW (regardless of reporting method). Experiment 3 found that reported depth of a wandering thought is perceived as shallow far more frequently than deep, regardless of report method. Results from Experiment 4 provided further evidence that many MW features (i.e., task relatedness, thought affect, movement) are highly similar across report type, although self-caught instances tend to be shorter in duration than probe-caught instances. Finally, the inclusion of both methods in Experiment 5 indicated no confound of using both methods concurrently, nor evidence that they are necessarily tapping into different constructs. A critical implication of this set of experiments is that self-caught reports captured a large number of MW instances, when compared to probe-caught methods. Although there is reluctance for MW researchers to use the self-caught method, these experiments demonstrate that self-caught reporting captures similar instances of MW. Aside from the more natural approach of self-catching, there may be cases where using the self-caught method is desirable (e.g., in tasks where probing an individual may be too disruptive to the task, such as driving). That said, the self-catching technique would benefit from further examination to understand whether individual differences in meta-awareness influence reporting ability, and how constant monitoring of attentional state influences task performance. Continued work in distinguishing qualitative differences in MW reports would build further support for the use of diverse methods.

Chapter 3 elaborated on the naturalness of task and/or settings to examine how MW rates vary when presentation style or task are manipulated. Results from Experiment 6 suggest that while MW is not influenced by lecture setting, retention performance is significantly better during live in-class viewing of lectures versus pre-recorded video lectures. This result may be due to students benefitting from having the lecture delivered by a professor who was physically present in the classroom. Interestingly, physical professor presence did not affect MW rates. Experiment 7 further investigated the impact of lecture setting on MW and retention by manipulating instructor visibility and social setting. Results indicate that neither MW nor retention were impacted by whether the instructor was visible in the video, but that MW was higher in the social setting condition wherein students watched the lecture as a group with a shared screen as compared to watching the lecture alone. Together, these two experiments offer a careful examination of how traditional classrooms differ from online learning environments, in terms of both cognitive impact and subjective experience. Finally, in Experiment 8 I pushed the envelope on what constitutes a "natural" task, moving away from lectures and reading material to a novel and fun activity: video game play. Furthermore, I examined whether the notion that "bigger is better" in regard to screen size applies to attention reports during an interactive leisure activity. Contrary to my hypothesis and related literature, none of MW rates or the other subjective ratings were influenced by screen size. Despite the lack of significant differences, the present experiment found that MW during a naturalistic and engaging video game task occurs at a rate of roughly 20% which is comparable to the lower bound reported in some simpler laboratory tasks.

Chapter 4 extended the research scope to include distraction as another form of inattention and examined whether MW and distraction occur at different rates between lab and real-world settings. In both experiments, I extended research on "natural" settings beyond a lab-based manipulation of "noise" by having participants complete a task outside the lab, a truly dynamic and natural environment. In Experiment 9, using a probe-caught paradigm, participants reported more distraction than MW in the natural setting but there were no differences in rates of inattention when inside the lab. When using a self-caught paradigm in Experiment 10, participants report more MW than distraction within the lab, but no differences in rates when outside of the lab. Of importance, these two experiments indicated that there may in fact be a dissociation in MW and distraction rates depending on whether the report is self-caught or probe-caught. Furthermore, retention performance was negatively impacted in the out of lab setting when probe-caught reporting was used, but not when self-caught reporting was used, suggesting that perhaps the opportunity to self-catch neutralizes the negative effect of MW on one's performance ability. Results from these experiments suggest that lab-based research on inattention may not generalize to dynamic real-world environments.

While these ten experiments provide a valuable contribution to MW research in natural tasks and settings, there is still much to be discovered about MW, including how MW can best be measured to reflect everyday thought patterns and how MW influences individual performance.

5.2 Limitations and future directions

In this concluding section, I will discuss the limitations that may have impacted my results and interpretations. First I will address the fact that many of my results, especially those comparing self-caught to probe-caught reporting, were in support of the null hypothesis (i.e., no difference). Next I will discuss three key points that this thesis did not address were: i) whether my interpretation of task and setting naturalness truly reflects real-world behaviour; ii) what the mechanisms behind self-caught and probe-caught reporting are; and iii) if there is stability in the impact of MW on retention performance. For each point I will make suggestions for future research that could help inform generalizability of the results and a better understanding of MW in natural tasks and settings.

5.2.1 A note on the significance of null hypotheses

A number of the experiments in the present experiment examined effects which have as of yet been underexplored (or absent) in the MW literature. While this suggests that the findings here provide novel and meaningful contributions to the field, the outcomes should (as with much research) be interpreted conservatively. Due to an inability to use prior literature to estimate effect sizes (e.g., when exploring potential differences between self-caught and probe-caught MW), my goal was to look for effects that were at least medium in size (i.e., Cohen's d equal to .5 for t-tests or .25 for ANOVAs). The reason for setting this goal was that smaller effects sizes may be more fragile, and therefore less replicable when additional real-world factors are introduced. For example, given that MW rates appear to be reliably influenced by interest and motivation, it is conceivable that the more "real world" an experiment becomes, the less likely it is for nuanced factors (such as display settings) to have a significant impact. Given that one of the overall goals of my thesis was to explore MW within natural tasks and settings, looking for robust and large effects was a critical point. As this decision could impact the ability to detect effects with regard to null hypothesis significance testing, I chose to also include Bayesian analyses which provides a measure of evidence *both* for and against a null hypothesis; traditional null hypothesis significance testing (NHST) only considers evidence *against* the null. The inclusion of Bayesian analyses provides an additional level of support to which outcomes are meaningful, either in support for or against the alternative hypothesis.

In the present thesis, the inclusion of Bayesian analyses allows me to report with some confidence that certain MW characteristics do not differ when using self-caught or probe-caught reporting. For example, intentionality of MW (BF_{01}: 3.82), thought affect (BF_{01}: 18.18), movement of MW thoughts (BF_{01}: 12.5), and emotional while MW (BF_{01}: 12.5), all provided moderate to strong support that the data fall under the same distribution curve. On the other hand, the Bayes analyses for a few characteristics (MW depth: BF_{10}: 1.04 and task relatedness: BF_{01}: .27) did not provide certainty with regard to differences or similarities between methods. These latter measures would warrant further research to better understand the situations under which self-caught or probe-caught reporting captures unique thought patterns.

5.2.2 Is this the real life? Is this just fantasy?

There are many compelling arguments on the importance of using natural settings and tasks in research, especially with regard to the ability to generalize between lab and everyday behaviour. In the present thesis I focused on using tasks that individuals (namely undergraduate students, as this was my target sample) would encounter in their everyday lives. Thus, all the tasks used (reading, audiobook listening, lecture watching, video game play) could be classified as “natural”. However, it is possible that not all the results obtained accurately represent “real-world” behaviour and one must consider whether the participant is interacting with the task in the same way that they would a task with real behavioural consequences. For example, from the perspective of a participant, the consequence of MW during an exam is likely quite different from the consequence of MW during an in-lab lecture task, and this may greatly impact participant engagement with the task. In my lab experiments, task performance would hold no real consequence, whereas task performance in the real classroom experiments demand attentional focus and the obligation to perform well. Although lab experiments could offer rewards for on-task performance, it is not clear whether the intrinsic motivation would be same as in everyday settings. Everyday MW may also be influenced differently than in the lab as individuals have greater choice and freedom to act in a way that suits their own needs and goals. For example, in a boring lab task there is little option for the participant to simply leave or move onto another task, whereas in everyday settings, individuals may shift their task to optimize their time and attention depending on the task requirements. As indicated above, there may be many overlapping factors that influence performance and the measurement of behaviours in the lab does not afford the same freedom, control or interactions as would occur in everyday settings. It is important then to consider whether any of the in-lab effects found in this thesis would generalise to naturalistic tasks within naturalistic settings. There is evidence that for lecture settings, the rate of MW and impact on performance may generalise across tasks: probe-caught rates of MW while watching lectures is relatively similar across in-lab lecture viewing tasks (range: 32-57%) and natural real-world lecture viewing (range: 40-49%). However, further work is needed to understand whether the quality of MW instances would be the same. For example, are intentional MW thoughts occurring at a similar rate between the lab and naturalistic everyday settings that afford more environmental fluctuations? As the present thesis only examined the nuances of MW characteristics within lab settings, this would be an important extension for future research.

One major strength of the present thesis is the use of self-reporting to measure MW as it represents a more natural way to capture these instances of inattention. However it is important to address that this method could be considered problematic due to the subjective nature of the reports. It could be argued that asking participants to self-report MW (either self-catch or probe-catch) presents a demand characteristic or priming into the task, thus influencing (perhaps inflating) reports. The subjective nature of self-reports of MW, and whether reports can be impacted by task characteristics or experimental instruction, is certainly a dilemma that the field of MW has yet to solve. On the other hand, the concern of subjectivity could easily be levelled at much of psychological measurement as a whole. That is, one might question whether asking people to report their attentional slips, subjective beliefs, or emotional states introduces demand characteristics that compromise their quantitative and qualitative measurements. Of course the role of experimental demand is always a factor that researchers need to be vigilant about, and combining self–reports with more objective measures of MW (e.g., eye tacking or fMRI) will help inform how information is being processed in the moments leading up to, as well as during, a MW episode.

Another limitation is that the majority of the natural tasks used did not vary greatly in user engagement (with the exception of gaming), thus my umbrella classification of tasks as natural versus unnatural does not address the variation that occurs within tasks. For example, watching a lecture, driving a vehicle, and reading a book vary greatly in terms of demands and cognitive resources each task requires. Although Experiment 2 (Chapter 2) hoped to examine learning tasks that vary in demand (i.e., reading a book versus watching a lecture), because the materials varied in content and no measure of task demand was collected, it is not yet clear how task demand, within natural tasks, may influence MW reports. That said, in Experiment 8 (Chapter 3) MW was reported, on average, 20% of the time during engaging video game play, a rate comparable to audiobook listening in Experiment 9 (Chapter 4). Thus, perhaps within experimental lab experiments (even those designed to be more naturalistic) MW will always have a lower-bound of roughly 20% as the tasks cannot reflect *truly* natural behaviour. Or perhaps MW rates of 20% are the lower-bound of MW frequency in either natural or unnatural settings.

Even when considering the desired engagement induced via playing a video game (Experiment 8, Chapter 3), it is possible this did not reflect real-world game play. For example, when playing a racing game, participants reported greater sense of presence and enjoyment when using a steering wheel controller as compared to a standard gaming controller (Williams, 2014). It is possible that the wheel was a more natural device for the given activity, and thus allowed more naturalistic movements. An experiment by Takatalo, Häkkinen, Kaistinen, and Nyman (2010) reported that playing a video game in the lab showed higher levels of attention and arousal (as per skin conductance, facial EMG, and heart rate) as compared to playing at home, suggesting both natural task and natural setting may also influence observed results.

One future direction that in-class testing could pursue would be the examination of learning analytics data to help support the reasons for why in-class and online lecture attention and performance may differ. In light of the results of Experiments 6 and 7 in Chapter 3, one obvious next step would be to monitor students' access of the online material and measure performance (e.g., retention) and subjective experience for both environments. Future more longitudinal investigations where lectures are systematically alternated between live and video presentation, as in blended-learning, would provide the opportunity to extend the present results and further determine the cumulative effects of professor and peer presence on student learning.

5.2.3 Who has a wandering mind? What are the mechanisms behind self-caught and probe-caught reporting?

Research by Schad et al. (2012) suggests that MW is not simply an all-or-none cognitive state, and that attention may decouple in a more graded manner. As highlighted in the introduction and Chapter 2, the subtyping of the MW experience may help sample qualitatively different parts of this gradient (e.g., intentional versus unintentional: Seli, Konishi, et al., 2018; Seli, Risko, Smilek, et al., 2016; Seli, Wammes, et al., 2015; Wammes, Boucher, et al., 2016). In some cases, the subcategorization of the MW report may influence outcomes. For example, Seli, Carriere, and Smilek (2015) report that spontaneous and deliberate MW are differentially related to aspects of mindfulness; and MW reported as unintentional was related to in-class quiz performance whereas unintentional MW was not (Wammes, Seli, et al., 2016). These nuances, and the distinctions found in the present thesis suggest that distinguishing between these alternatives may be a fruitful avenue for future research in terms of determining whether a specific *type* of MW is driving the effects. It could also be informative to carefully examine covariates that might influence MW reports and the relationship between measurement methodologies, e.g., time on task (Farley et al., 2013; Thomson, Seli, Besner, & Smilek, 2014; Wammes, Boucher, et al. 2016), task difficulty (Feng, D'Mello, & Graesser, 2013), or probe frequency (Seli et al., 2013). Of particular relevance to this thesis was whether task features, such as their interest value, influences MW reports. A key goal of the present thesis was to use tasks that would be considered more natural (i.e., reading) than many other commonly used paradigms (e.g., the SART). While the tasks used were more natural, the extent to which participants were interested in engaging with the task may have influenced outcomes. The results of the video game and MW experiment (Experiment 8, Chapter 3) make a reasonable case for why task characteristics should be considered more carefully in attention paradigms. That individuals were not completely immersed in this "fun" task, as well as the fact that video game influenced MW rates, suggests that what researchers consider natural or engaging may fluctuate quite a bit based on individual preference. Although all the tasks in the present this were chosen to be natural, there would likely be great variation in MW rates between a competitive video game player and a non-gamer, or between two natural tasks within a lab setting (e.g., as was seen for reading versus lecture watching in Experiment 2, Chapter 2). Future research would benefit from a careful manipulation of task features to gain a better understanding of what leads to MW during natural tasks, both with regard to task features and how they combine with individual preferences.

In addition to an examination of how task features or experimental instruction influences MW rates, an open question for future investigation is how MW rates and the task manipulations that influence them vary as a function of individual differences, such as aptitude and learning motivation styles. For example, Wilson and Korn (2007) state that individual differences (i.e., motivation and WMC) are more influential to fluctuations in attention than task duration. Their research suggests that the relationships between MW and task are complex and not necessarily linear. Imagine the following scenario: a student with low WMC is attending a boring lecture, but this student has high motivation due to the need to perform well on an exam. If a researcher simply examines how much MW occurs in this scenario without considering the additional factors at play (as most MW experiments do), it would be unreasonable to make generalizations of MW rates in this setting. In the present experiments, I do not consider any individual differences as potential mediating influences on MW rates, but in everyday activities multiple factors most certainly will be present and most likely will influence MW rates concurrently. The framework of ecological validity promotes the practice of looking for effects in controlled lab settings and then determining whether these generalize to in natural everyday tasks and settings (and vice-versa). For example, if a researcher found that MW rates in the lab were higher in for individuals reporting low WMC, the relationship should then be tested in an everyday environment to determine the robustness of the effect when task setting is highly variable. On the other hand, for high MW rates that are reported within an everyday classroom setting, researchers would benefit from examining each element of that environment to determine which specific factors are contributing to these rates. Applying the cognitive ethology framework to future MW research is one way to further examine why MW rates may differ across tasks and settings varying in naturalness.

One important individual difference that would be valuable to examine in relation to the differences between self-caught and probe-caught MW reports concerns whether participants have greater difficulty noticing that their minds have wandered within certain tasks, with one possibility being that task passivity leads to reduced meta-awareness. Moving forward, assessing individual rates of meta-awareness or mindfulness could determine whether meta-awareness influences the likelihood to capture one specific type of MW over the other (i.e., intentional versus unintentional). van Vugt & Broers (2016) suggest that those with better mindfulness will have greater thought awareness and thus reduced tendency to get "stuck" in MW instances, this is a hypothesis that could be tested with both self- and probe-caught methods.

5.2.4 The stability of the impact of MW on retention performance. Is there a cost associated with MW?

Prior research using probe-caught methods, but not self-caught methods, has found that MW reports are often negatively correlated with task performance (Franklin, Smallwood, & Schooler, 2011; Lindquist & Mclean, 2011; Schooler et al., 2004; Smallwood et al., 2008b; Szpunar, Moulton, et al., 2013; Varao Sousa et al., 2013). However, this relationship has not been found for self-caught reports, nor for a number of probe-caught MW experiments (Sayette et al., 2009; Schooler et al., 2004; Varao-Sousa & Kingstone, 2018; Wammes & Smilek, 2017:

MW and retention relationship was moderated by retention duration and MW type). In a nuanced examination of the impact of MW on performance, Jing, Szpunar and Schacter (2016) found that thoughts related to a lecture do not negatively impact test performance whereas unrelated thoughts negatively predict memory test performance.

In the present thesis, only four of the 10 experiments were substantially powered to examine the correlation between MW and task performance and only half these experiments found a significant relationship. In Experiment 4 (Chapter 2), self-caught reports of MW were not related to task performance whereas increased probe-caught reports were related to poorer task performance. In contrast, neither self-caught nor probe-caught MW rates were related to retention performance in the live classroom setting of Experiment 5 (Chapter 2). Experiments 6 and 7 in Chapter 3 saw significant negative relationships, however for Experiment 6 only the video recorded version of the lecture had a significant relationship, not the live version. Taken together, it appears that MW negatively impacts performance during video recordings of lectures, but not live versions of lectures. One reason this might be the case is that task motivation may be different between these two settings.

Future work would certainly benefit from investigating whether specific “types” of MW drive relationships between MW and retention test performance. Within a probe-caught paradigm Seli, Cheyne, et al., (2015) report that both unintentional and intentional MW are predictive of declines in performance, but that the overall relationship may be driven by task motivation. Replicating this finding with self-caught MW and within naturalistic tasks would be valuable. Seli et al. (2014) suggest that the resources required for maintaining task focus “might vary as a function of depth of mind wandering [and] … one would expect deficits in task-related performance only if the requirements of a task exceed the amount of resources directed toward that task" (p. 4). Seli, Carriere et al. (2018) and Thomson et al. (2013) also find evidence that individuals can MW or media multitask without performance costs in situations where task demands are low or resource conflicts are minimal. In line with the results in Chapter 2 - Experiment 3, this may suggest that shallow MW during a video lecture, which requires low resources, may not impact content retention. On the other hand, if a task only *appears* to require low resources then any type of MW would negatively impact retention.

Distractions, such as other people talking, construction and noisy environments have also been found to negatively impact task performance (Blasiman, Larabee, & Fabry, 2018; Stawarczyk, Majerus, Maj, Van der Linden, & D'Argembeau, 2011; Zeamer & Fox Tree, 2013). In light of the evidence from Chapter 4, that MW and distraction occur at different frequencies inside and outside of the lab, it would be informative to distinguish the situations in which task performance suffers as a result of MW or distraction.

5.2.5 What is MW? How we define it and measure it matters

The study of MW within the field of cognitive psychology research is relatively new, and researchers have come to no consensus on how to define the term or individuals' experience of the occurrence (Callard, Smallwood, Golchert, & Margulies, 2013; Mills, Raffaelli, Irving, Stan, & Christoff, 2018; Seli, Kane, et al., 2018; Seli, Ralph, et al., 2017; Weinstein, 2017). This lack of unity in the field means that researchers often provide participants with divergent (and perhaps conflicting) definitions. Within the present thesis, MW was consistently defined as thinking about something unrelated to the assigned task, thus results found across experiments should be consistently tapping into the same cognitive experience. However, when comparing the present results to other experiments it could be useful to consider how various definitions of MW may impact results given that the inconsistent definition and presentation of MW prompts has been shown to have an impact on experimental outcomes. For example, Weinstein, De Lima, and van der Zee (2018) manipulated whether participants were asked to report being “on task” or to report “mind wandering”. The results indicated that data that would be considered as MW were reported more often under the “mind wandering” framed question than the “on task” question. The authors speculate that this relatively small difference in prompt presentation may obscure existing data interpretations. In another examination of the influence of MW probes on MW rates, Seli, Carriere, Levene, and Smilek (2013) examined whether probe frequency would influence MW rates. While completing a sustained attention task, MW was reported more often when there was a greater lag between probes, as compared to probes that occurred more frequently. These are just two examples of how measurement tools may be influencing MW reports, above and beyond variability due to task, setting, or report method. A consistent and explicit definition of MW used by all researchers studying the cognitive state would help in determining whether results can be replicated across tasks and settings.

5.3 Conclusion

Overall, the results of the present thesis suggest that individuals frequently report MW, regardless of whether reporting is probe-caught or self-caught. The impact of inattention and the when and why our minds wander is a question that interests many, as it is such a prevalent occurrence in everyday behaviour. The work presented in this thesis highlights the critical importance of conducting experiments in contexts where MW occurs in naturalistic tasks and settings. Put simply, the generalizability of MW results to settings and tasks outside simplistic lab paradigms is as yet poorly understood, and ripe for future study. Indeed, it seems that the less used but more ecologically valid method of self-caught reports capture MW at a high rate, which often mimics the nuances of probe-caught reporting. That said, some importance differences did emerge between the two methods, and researchers would benefit from a careful examination of each when determining which to use in their own design. Furthermore, this research provided evidence that inattention rates may be influenced by setting naturalness, a conclusion that has great implications for theories of human attention, as the majority of attention research occurs within ecologically challenged lab settings. Continuing the investigation of MW across paradigms varying in naturalness will enable better assessment of the impact that MW has on task performance, which will in turn inform theories of human (in)attention and performance.

References

Abdous, M., & Yoshimura, M. (2010). Learner outcomes and satisfaction: A comparison of live video-streamed instruction, satellite broadcast instruction, and face-to-face instruction. *Computers & Education*, *55*(2), 733–741. https://doi.org/10.1016/j.compedu.2010.03.006

Anderson, G., & Brown, R. I. F. (1984). Real and laboratory gambling, sensation-seeking and arousal. *British Journal of Psychology*, *75*(3), 401–410. https://doi.org/10.1111/j.2044-8295.1984.tb01910.x

Antrobus, J. S., Singer, J. L., & Greenberg, S. (1966). Studies in the stream of consciousness : experimental enhancement and suppression of spontaneous cognitive processes. *Perceptual and Motor Skills*, *23*, 399–417. https://doi.org/10.2466/pms.1966.23.2.399

Aragon, S. (2010). Creating social presence in online environments. In *New Directions for Adult and Continuing Education (…*. San Francisco: Jossey Bass.

Baddeley, A. (1992). Working Memory: The Interface between Memory and Cognition. *Journal of Cognitive Neuroscience*, *4*(3), 281–288. https://doi.org/10.1162/jocn.1992.4.3.281

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition*, *20*(4), 1604–1611. https://doi.org/10.1016/j.concog.2011.08.007

Baldwin, C. L., Roberts, D. M., Barragan, D., Lee, J. D., Lerner, N., & Higgins, J. S. (2017). Detecting and Quantifying Mind Wandering during Simulated Driving. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00406

Banaji, M. (1989). The bankruptcy of everyday memory. *American Psychologist*, *44*(9), 1185–1193. Retrieved from http://psycnet.apa.org/record/1990-00387-001

Banaji, M., & Crowder, R. (1991). Some everyday thoughts on ecologically valid methods.

American Psychologist, *46*(1), 78–79. Retrieved from http://psycnet.apa.org/record/1991-11752-001

Banbury, S., & Berry, D. C. (1998). Disruption of office-related tasks by speech and office noise. *British Journal of Psychology*, *89*(3), 499–517. https://doi.org/10.1111/j.2044-8295.1998.tb02699.x

Banks, J. B., & Boals, A. (2017). Understanding the role of mind wandering in stress-related working memory impairments. *Cognition and Emotion*, *31*(5), 1023–1030. https://doi.org/10.1080/02699931.2016.1179174

Banks, J. B., Welhaf, M. S., Hood, A. V. B., Boals, A., & Tartar, J. L. (2016). Examining the role of emotional valence of mind wandering: All mind wandering is not equal. *Consciousness and Cognition*, *43*, 167–176. https://doi.org/10.1016/J.CONCOG.2016.06.003

Barker, R. (1968). *Ecological psychology: Concepts and methods for studying the environment of human behavior*. Stanford University Press.

Barkley, R. A. (1991). The ecological validity of laboratory and analogue assessment methods of ADHD symptoms. *Journal of Abnormal Child Psychology*, *19*(2), 149–178. https://doi.org/10.1007/BF00909976

Bernard, R. M., Abrami, P. C., Lou, Y., Borokhovski, E., Wade, A., Wozney, L., … Huang, B. (2004). How Does Distance Education Compare With Classroom Instruction? A Meta-Analysis of the Empirical Literature. *Review of Educational Research*, *74*(3), 379–439. https://doi.org/10.3102/00346543074003379

Berthié, G., Lemercier, C., Paubel, P.-V., Cour, M., Fort, A., Galéra, C., … Maury, B. (2015). The restless mind while driving: drivers’ thoughts behind the wheel. *Accident; Analysis and*

Prevention, *76*, 159–165. https://doi.org/10.1016/j.aap.2015.01.005

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., … Devins, G. (2006). Mindfulness: A Proposed Operational Definition. *Clinical Psychology: Science and Practice*, *11*(3), 230–241. https://doi.org/10.1093/clipsy.bph077

Bond, C. F., & Titus, L. J. (1983). Social facilitation: A meta-analysis of 241 studies. *Psychological Bulletin*, *94.2*, 265.

Bracken, C. C., Pettey, G., Guha, T., & Rubenking, B. E. (2010). Sounding Out Small Screens and Telepresence. *Journal of Media Psychology*, *22*(3), 125–137. https://doi.org/10.1027/1864-1105/a000017

Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*, *40*(4), 509–513. https://doi.org/10.1152/advan.00109.2016

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513. Retrieved from http://psycnet.apa.org/record/1978-06857-001

Bryson, B., & Roberts, W. (2003). *A short history of nearly everything*. Retrieved from https://europepmc.org/articles/pmc2794571

Burgess, P., Alderman, N., Evans, J., Emslie, H., & Wilson, B. (1998). The ecological validity of tests of executive function. *Journal Ofthe International Neuropsychological Society*, *4*, 547–558.

Burns, C. G., & Fairclough, S. H. (2015). Use of auditory event-related potentials to measure immersion during a computer game. *International Journal of Human-Computer Studies*, *73*, 107–114. https://doi.org/10.1016/J.IJHCS.2014.09.002

Callard, F., Smallwood, J., Golchert, J., & Margulies, D. S. (2013). The era of the wandering mind? Twenty-first century research on self-generated mental activity. *Frontiers in Psychology*, *4*(DEC), 891. https://doi.org/10.3389/fpsyg.2013.00891

Cameron, P., & Giuntoli, D. (1972). Consciousness Sampling in the College Classroom or Is Anybody Listening?. *Intellect*, *101*(2343), 63–64. Retrieved from https://eric.ed.gov/?id=EJ065971

Canadian Virtual University. (2012). *Online University Education in Canada : Challenges and Opportunities*. (January).

Cao, A., Chintamani, K. K., Pandya, A. K., & Ellis, R. D. (2009). NASA TLX: Software for assessing subjective mental workload. *Behavior Research Methods*, *41*(1), 113–117. https://doi.org/10.3758/BRM.41.1.113

Carriere, J. S. A., Cheyne, J. A., & Smilek, D. (2008). Everyday attention lapses and memory failures: The affective consequences of mindlessness. *Consciousness and Cognition*, *17*(3), 835–847. https://doi.org/10.1016/j.concog.2007.04.008

Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, *80*, 108–121. https://doi.org/10.1016/j.compedu.2014.08.015

Cheyne, J. A., Solman, G. J. F., Carriere, J. S. A., & Smilek, D. (2009). Anatomy of an error: a bidirectional state model of task engagement/disengagement and attention-related errors. *Cognition*, *111*(1), 98–113. https://doi.org/10.1016/j.cognition.2008.12.009

Chisholm, J. D., & Kingstone, A. (2015). Action video games and improved attentional control: Disentangling selection- and response-based processes. *Psychonomic Bulletin & Review*, *22*(5), 1430–1436. https://doi.org/10.3758/s13423-015-0818-3

Christoff, K., Gordon, A. M., Smallwood, J., Smith, R., & Schooler, J. W. (2009). Experience sampling during fMRI reveals default network and executive system contributions to mind wandering. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(21), 8719–8724. https://doi.org/10.1073/pnas.0900234106

Colavita, F. B. (1974). Human sensory dominance. *Perception & Psychophysics*, *16*(2), 409–412. https://doi.org/10.3758/BF03203962

Cowley, J. A. (2013). Off Task Thinking Types and Performance Decrements During Simulated Automobile Driving; Off Task Thinking Types and Performance Decrements During Simulated Automobile Driving. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1214–1218. https://doi.org/10.1177/1541931213571270

Crook, C., & Schofield, L. (2017). The video lecture. *Internet and Higher Education*, *34*, 56–64. https://doi.org/10.1016/j.iheduc.2017.05.003

Csikszentmihalyi, M., & Larson, R. (2014). Validity and Reliability of the Experience-Sampling Method. In *Flow and the Foundations of Positive Psychology* (pp. 35–54). https://doi.org/10.1007/978-94-017-9088-8_3

Csikszentmihalyi, M., Larson, R., & Prescott, S. (1977). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence*, *6*(3), 281–294. https://doi.org/10.1007/BF02138940

Curci, A., Lanciano, T., Soleti, E., & Rimé, B. (2013). Negative emotional experiences arouse rumination and affect working memory capacity. *Emotion*, *13*(5), 867–880. https://doi.org/10.1037/a0032492

d'Alessio, M. A., Lundquist, L. L., Schwartz, J. J., Pedone, V., Pavia, J., & Fleck, J. (2019). Social presence enhances student performance in an online geology course but depends on instructor facilitation. *Journal of Geoscience Education*, *67*(3), 222–236.

https://doi.org/10.1080/10899995.2019.1580179

De Cesarei, A., & Codispoti, M. (2006). When does size not matter? Effects of stimulus size on affective modulation. *Psychophysiology*, *43*(2), 207–215. https://doi.org/10.1111/j.1469-8986.2006.00392.x

Demetriadis, S., & Pombortsis, A. (2007). *e-Lectures for Flexible Learning : a Study on their Learning Efficiency*. *10*, 147–157.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., … Horrey, W. J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *PNAS*, *113*(10), 2636–2641. Retrieved from http://www.pnas.org/content/113/10/2636.full.pdf

Dixon, P., & Li, H. (2013). Mind wandering in text comprehension under dual-task conditions. *Frontiers in Psychology*, *4*(OCT), 1–14. https://doi.org/10.3389/fpsyg.2013.00682

Eastwood, J. D., Frischen, A., Fenske, M. J., & Smilek, D. (2012). The Unengaged Mind. *Perspectives on Psychological Science*, *7*(5), 482–495. https://doi.org/10.1177/1745691612456044

Entertainment software association. (n.d.). U.S. Video Game Industry Generates $30.4 Billion in Revenue for 2016 - The Entertainment Software Association. Retrieved October 18, 2018, from http://www.theesa.com/article/u-s-video-game-industry-generates-30-4-billion-revenue-2016/

Erichsen, E. A., & Bolliger, D. U. (2010). Towards understanding international graduate student isolation in traditional and online environments. *Educational Technology Research and Development*, *59*(3), 309–326. https://doi.org/10.1007/s11423-010-9161-6

Farley, J., Risko, E. F., & Kingstone, A. (2013). Everyday attention and lecture retention: the effects of time, fidgeting, and mind wandering. *Frontiers in Psychology*, *4*(September), 619. https://doi.org/10.3389/fpsyg.2013.00619

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fawcett, J., Risko, E., & Kingstone, A. (2015). *The Handbook of Attention*. Retrieved from https://books.google.ca/books?hl=en&lr=&id=z8XlCgAAQBAJ&oi=fnd&pg=PR7&ots=4vkbKzYq2F&sig=ZgEn-ODtgGqfDHRE9Og0_kwFdrU

Feng, S., D'Mello, S., & Graesser, A. C. (2013). Mind wandering while reading easy and difficult texts. *Psychonomic Bulletin & Review*, *20*(3), 586–592. https://doi.org/10.3758/s13423-012-0367-y

Ferguson, J., & Tryjankowski, A. M. (2009). Online versus face-to-face learning: looking at modes of instruction in Master's-level courses. *Journal of Further and Higher Education*, *33*(3), 219–228. https://doi.org/10.1080/03098770903026149

Forster, S., & Lavie, N. (2007). High Perceptual Load Makes Everybody Equal. *Psychological Science*, *18*(5), 377–381. https://doi.org/10.1111/j.1467-9280.2007.01908.x

Forster, S., & Lavie, N. (2009). Harnessing the wandering mind: The role of perceptual load. *Cognition*, *111*(3), 345–355. https://doi.org/10.1016/j.cognition.2009.02.006

Forster, S., & Lavie, N. (2014). Distracted by your mind? Individual differences in distractibility predict mind wandering. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *40*(1), 251–260. https://doi.org/10.1037/a0034108

Foulsham, T, Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision Research*, *51*(17), 1920–1931. Retrieved from https://www.sciencedirect.com/science/article/pii/S0042698911002392

Foulsham, Tom, Chapman, C., Nasiopoulos, E., & Kingstone, A. (2014). Top-down and bottom-up aspects of active search in a real world environment. *Canadian Journal of Experimental Psychology*, *68*(1), 8–19. https://doi.org/10.1037/cep0000004

Foulsham, Tom, Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: effects of social status and speaking. *Cognition*, *117*(3), 319–331. https://doi.org/10.1016/j.cognition.2010.09.003

Foulsham, Tom, Farley, J., & Kingstone, A. (2013). Mind wandering in sentence reading: decoupling the link between mind and eye. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, *67*(1), 51–59. https://doi.org/10.1037/a0030217

Franklin, M. S., Mooneyham, B. W., Baird, B., & Schooler, J. W. (2014). Thinking one thing, saying another: The behavioral correlates of mind-wandering while reading aloud. *Psychonomic Bulletin & Review*, *21*(1), 205–210. https://doi.org/10.3758/s13423-013-0468-2

Franklin, M. S., Mrazek, M. D., Anderson, C. L., Smallwood, J., Kingstone, A., & Schooler, J. W. (2013). The silver lining of a mind in the clouds: Interesting musings are associated with positive mood while mind-wandering. *Frontiers in Psychology*, *4*(AUG), 583. https://doi.org/10.3389/fpsyg.2013.00583

Franklin, M. S., Smallwood, J., & Schooler, J. W. (2011). Catching the mind in flight: using behavioral indices to detect mindless reading in real time. *Psychonomic Bulletin & Review*, *18*(5), 992–997. https://doi.org/10.3758/s13423-011-0109-6

Franklin, M. S., Smallwood, J., Zedelius, C. M., Broadway, J. M., & Schooler, J. W. (2015).

Unaware yet reliant on attention: Experience sampling reveals that mind-wandering impedes implicit learning. *Psychonomic Bulletin & Review*, *23*(1), 223–229. https://doi.org/10.3758/s13423-015-0885-5

Galéra, C., Orriols, L., M'Bailara, K., Laborey, M., Contrand, B., Ribéreau-Gayon, R., … Lagarde, E. Mind wandering and driving: responsibility case-control study. , 345 BMJ: British Medical Journal § (2012).

Garrison, D. R. (2011). *E-Learning in the 21st Century: A Framework for Research and Practice*. Retrieved from https://books.google.com/books?hl=en&lr=&id=aodjWyjxYbYC&pgis=1

Giambra, L. M., & Grodsky, A. (1989). Task-Unrelated Images and Thoughts While Reading. In J. E. Shorr, P. Robin, J. A. Connella, & M. Wolpin (Eds.), *Imagery: Current Perspective*. https://doi.org/10.1007/978-1-4899-0876-6

Grainger, B. (2013). *University of London International Programmes: MOOC Report*. Retrieved from http://www.londoninternational.ac.uk/sites/default/files/documents/mooc_report-2013.pdf

Grodsky, A., & Giambra, L. M. (1990). The Consistency Across Vigilance and Reading Tasks of Individual Differences in the Occurrence of Task-Unrelated and Task-Related Images and Thoughts. *Imagination, Cognition and Personality*, *10*(1), 39–52. https://doi.org/10.2190/6QG5-CXVV-4XUR-7P3K

Ham, T. (2004). Play Fast And Loose (washingtonpost.com). Retrieved October 30, 2018, from Washington Post website: http://www.washingtonpost.com/wp-dyn/articles/A60547-2004Jul18.html

Hawkins, G., Mittner, M., Boekel, W., & Heathcote, A. (2015). Toward a model-based cognitive neuroscience of mind wandering. *Neuroscience*, *310*, 290–305. Retrieved from https://www.sciencedirect.com/science/article/pii/S0306452215008805

He, J., Becic, E., Lee, Y.-C., & McCarley, J. S. (2011). Mind wandering behind the wheel: performance and oculomotor correlates. *Human Factors*, *53*(1), 13–21. https://doi.org/10.1177/0018720810391530

Heo, M., & Leon, A. C. (2010). Sample Sizes Required to Detect Two-Way and Three-Way Interactions Involving Slope Differences in Mixed-Effects Linear Models. *Journal of Biopharmaceutical Statistics*, *20*(4), 787–802. https://doi.org/10.1080/10543401003618819

Hines, A. M., & Shaw, G. A. (2013). Intrusive thoughts, sensation seeking, and drug use in college students. *Bulletin of the Psychonomic Society*, *31*(6), 541–544. https://doi.org/10.3758/BF03337347

Hollis, R. Ben, & Was, C. A. (2009). Mind Wandering and Online Learning: How Working Memory, Interest, and Mind Wandering Impact Learning from Videos. *Roceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36), 3084–3089.

Hollis, R. B., & Was, C. A. (2016). Mind wandering, control failures, and social media distractions in online learning. *Learning and Instruction*, *42*, 104–112. https://doi.org/10.1016/J.LEARNINSTRUC.2016.01.007

Hou, J., Nam, Y., Peng, W., & Lee, K. M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior*, *28*(2), 617–623. https://doi.org/10.1016/j.chb.2011.11.007

Huguet, P., Galvaing, M. P., Monteil, J. M., & Dumas, F. (1999). Social presence effects in the Stroop task: Further evidence for an attentional view of social facilitation. *Journal of Personality and Social Psychology*, *77*, 1011.

Ivory, J. D., & Magee, R. G. (2009). You Can't Take It with You? Effects of Handheld Portable Media Consoles on Physiological and Psychological Responses to Video Game and Movie Content. *CyberPsychology & Behavior*, *12*(3), 291–297. https://doi.org/10.1089/cpb.2008.0279

Jackson, J. D., & Balota, D. A. (2012). Mind-wandering in younger and older adults: Converging evidence from the sustained attention to response task and reading for comprehension. *Psychology and Aging*, *27*(1), 106–119. https://doi.org/10.1037/a0023933

Joksimović, S., Gašević, D., Kovanović, V., Riecke, B. E., & Hatala, M. (2015). Social presence in online discussions as a process predictor of academic performance. *Journal of Computer Assisted Learning*, *31*(6), 638–654. https://doi.org/10.1111/jcal.12107

Joksimović, Srećko, Poquet, O., Kovanović, V., Dowell, N., Mills, C., Gašević, D., … Brooks, C. (2018). How Do We Model Learning at Scale? A Systematic Review of Research on MOOCs. *Review of Educational Research*, *88*(1), 43–86. https://doi.org/10.3102/0034654317740335

Jordan, K. (2015). Massive open online course completion rates revisited: Assessment, length and attrition. *International Review of Research in Open and Distance Learning*, *16*(3), 341–358. https://doi.org/10.19173/irrodl.v16i3.2112

Kam, J. W. Y., & Handy, T. C. (2013). The neurocognitive consequences of the wandering mind: a mechanistic account of sensory-motor decoupling. *Frontiers in Psychology*, *4*(October), 725. https://doi.org/10.3389/fpsyg.2013.00725

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007a). For Whom the Mind Wanders, and When. *Psychological Science*, *18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007b). For Whom the Mind Wanders, and When. *Psychological Science*, *18*(7), 614–621. https://doi.org/10.1111/j.1467-9280.2007.01948.x

Kane, M. J., Gross, G. M., Chun, C. A., Smeekens, B. A. B., Meier, M. E., Silvia, P. J., … Kwapil, T. R. (2017). For Whom the Mind Wanders, and When, Varies Across Laboratory and Daily-Life Settings. *Psychological Science*, *28*(9), 095679761770608. https://doi.org/10.1177/0956797617706086

Kane, M. J., Gross, G. M., Chun, C. A., Smeekens, B. A., Meier, M. E., Silvia, P. J., & Kwapil, T. R. (2017). For Whom the Mind Wanders, and When, Varies Across Laboratory and Daily-Life Settings. *Psychological Science, 28*(9), 1271–1289. https://doi.org/10.1177/0956797617706086

Khalil, H., & Ebner, M. (2014). MOOCs Completion Rates and Possible Methods to Improve Retention - A Literature Review. *EdMedia: World Conference on Educational Media and Technology*, *2014*(1), 1305–1313. Retrieved from https://www.learntechlib.org/p/147656/

Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science (New York, N.Y.)*, *330*(6006), 932. https://doi.org/10.1126/science.1192439

Kim, D., & Kim, D. J. (2012). Effect of screen size on multimedia vocabulary learning. *British Journal of Educational Technology*, *43*(1), 62–70. https://doi.org/10.1111/j.1467-8535.2010.01145.x

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks inonline lecture videos. *Proceedings of the First ACM Conference on Learning @ Scale Conference - L@S '14*, 31–40. https://doi.org/10.1145/2556325.2566237

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive Ethology: a new approach for studying human cognition. *British Journal of Psychology (London, England : 1953)*, *99*(Pt 3), 317–340. https://doi.org/10.1348/000712607X251243

Kingstone, A., Smilek, D., Ristic, J., Kelland Friesen, C., & Eastwood, J. D. (2003). Attention, Researchers! It Is Time to Take a Look at the Real World. *Current Directions in Psychological Science*, *12*(5), 176–180. https://doi.org/10.1111/1467-8721.01255

Kizilcec, R., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: effects on information retention, visual attention, and affect. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Retrieved from https://dl.acm.org/citation.cfm?id=2557207

Kopp, K., & D'Mello, S. (2015). The Impact of Modality on Mind Wandering during Comprehension. *Applied Cognitive Psychology*, *30*(1), 29–40. https://doi.org/10.1002/acp.3163

Kovanović, V., Joksimović, S., Poquet, O., Hennis, T., Čukić, I., de Vries, P., … Gašević, D. (2018). Exploring communities of inquiry in Massive Open Online Courses. *Computers and Education*, *119*, 44–58. https://doi.org/10.1016/j.compedu.2017.11.010

Kuehner, C., Welz, A., Reinhard, I., Alpers, G. W., Pine, D., & Quinn, K. (2017). Lab meets real life: A laboratory assessment of spontaneous thought and its ecological validity. *PLOS ONE*, *12*(9), e0184488. https://doi.org/10.1371/journal.pone.0184488

Laidlaw, K. E. W., Foulsham, T., Kuhn, G., & Kingstone, A. (2011). Potential social interactions are important to social attention. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(14), 5548–5553. https://doi.org/10.1073/pnas.1017022108

Langton, S., Watt, R., & Bruce, I. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50–59. Retrieved from https://www.sciencedirect.com/science/article/pii/S1364661399014369

Levinson, D. B., Smallwood, J., & Davidson, R. J. (2012). The Persistence of Thought. *Psychological Science*, *23*(4), 375–380. https://doi.org/10.1177/0956797611431465

Lindquist, S. I., & McLean, J. P. (2011). Daydreaming and its correlates in an educational environment. *Learning and Individual Differences*, *21*(2), 158–167. https://doi.org/10.1016/j.lindif.2010.12.006

Masicampo, E. J., & Baumeister, R. F. (2011). Consider it done! Plan making can eliminate the cognitive effects of unfulfilled goals. *Journal of Personality and Social Psychology*, *101*(4), 667–683. https://doi.org/10.1037/a0024192

Masson, M. E. J. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behavior Research Methods*, *43*(3), 679–690. https://doi.org/10.3758/s13428-010-0049-5

McKinney, D., Dyck, J. L., & Luber, E. S. (2009). iTunes University and the classroom: Can podcasts replace Professors? *Computers & Education*, *52*(3), 617–623. https://doi.org/10.1016/j.compedu.2008.11.004

Mcvay, J. C., & Kane, M. J. (2010). Does mind wandering reflect executive function or executive failure? Comment on Smallwood and Schooler (2006) and Watkins (2008). *Psychological Bulletin*, *136*(136), 188–197.

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(1), 196–204.

McVay, J. C., & Kane, M. J. (2013). Dispatching the wandering mind? Toward a laboratory method for cuing “spontaneous” off-task thought. *Frontiers in Psychology*, *4*(SEP), 0–16. https://doi.org/10.3389/fpsyg.2013.00570

McVay, J. C., Kane, M. J., & Kwapil, T. R. (2009). Tracking the train of thought from the laboratory into everyday life: An experience-sampling study of mind wandering across controlled and ecological contexts. *Psychonomic Bulletin & Review*, *16*(5), 857–863. https://doi.org/10.3758/PBR.16.5.857

McVay, J. C., Unsworth, N., McMillan, B. D., & Kane, M. J. (2013). Working memory capacity does not always support future-oriented mind-wandering. *Canadian Journal of Experimental Psychology = Revue Canadienne de Psychologie Expérimentale*, *67*(1), 41–50. https://doi.org/10.1037/a0031252

Meyer, R. (2012). What It’s Like to Teach a MOOC (and What the Heck’s a MOOC?) - The Atlantic. Retrieved July 5, 2015, from http://www.theatlantic.com/technology/archive/2012/07/what-its-like-to-teach-a-mooc-and-what-the-hecks-a-mooc/260000/

Mills, C., Raffaelli, Q., Irving, Z. C., Stan, D., & Christoff, K. (2018). Is an off-task mind a freely-moving mind? Examining the relationship between different dimensions of thought. *Consciousness and Cognition*, *58*, 20–33. https://doi.org/10.1016/J.CONCOG.2017.10.003

Mook, D. G. (1984). In defense of external invalidity. *American Psychologist*, *38*(4), 229–232.

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *67*(1), 11–18. https://doi.org/10.1037/a0031569

Morrison, A. B., Goolsarran, M., Rogers, S. L., & Jha, A. P. (2014). Taming a wandering attention: short-form mindfulness training in student cohorts. *Frontiers in Human*

Neuroscience, *7*. https://doi.org/10.3389/fnhum.2013.00897

Mrazek, M. D., Smallwood, J., Franklin, M. S., Chin, J. M., Baird, B., & Schooler, J. W. (2012). The Role of Mind-Wandering in Measurements of General Aptitude. *Journal of Experimental Psychology: General*, *141*(4), 788–798. https://doi.org/10.1037/a0027968

Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2012). Mindfulness and mind-wandering: Finding convergence through opposing constructs. *Emotion*, *12*(3), 442–448. https://doi.org/10.1037/a0026678

O'Neil, C. (2016). *Weapons of math destruction : how big data increases inequality and threatens democracy*. Retrieved from https://books.google.ca/books?hl=en&lr=&id=cbwvDwAAQBAJ&oi=fnd&pg=PA1&dq=o%27neil+weapons&ots=Xjv8UYIvXd&sig=5NUaBWlduYBFOEEFrU9DDFtqTbI#v=onepage&q=o%27neil weapons&f=false

Olivers, C. N. L., & Nieuwenhuis, S. (2005). The Beneficial Effect of Concurrent Task-Irrelevant Mental Activity on Temporal Attention. *Psychological Science*, *16*(4), 265–269. https://doi.org/10.1111/j.0956-7976.2005.01526.x

Onah, D. F. O., Sinclair, J., & Boyatt, R. (2014, January 8). Dropout rates of massive open online courses : behavioural patterns. *EDULEARN14 Proceedings*. Retrieved from http://wrap.warwick.ac.uk/65543/1/WRAP_9770711-cs-070115-edulearn2014.pdf

Ottaviani, C., & Couyoumdjian, A. (2013). Pros and cons of a wandering mind: A prospective study. *Frontiers in Psychology*, *4*(AUG), 1–9. https://doi.org/10.3389/fpsyg.2013.00524

Ottaviani, C., Shahabi, L., Tarvainen, M., Cook, I., Abrams, M., & Shapiro, D. (2015). Cognitive, behavioral, and autonomic correlates of mind wandering and perseverative cognition in major depression. *Frontiers in Neuroscience*, *8*, 433.

https://doi.org/10.3389/fnins.2014.00433

Ottaviani, C., Shapiro, D., & Couyoumdjian, A. (2013). Flexibility as the key for somatic health: From mind wandering to perseverative cognition. *Biological Psychology*, *94*(1), 38–43. https://doi.org/10.1016/J.BIOPSYCHO.2013.05.003

Peng, A., Kirkham, N. Z., & Mareschal, D. (2018). Information processes of task-switching and modality-shifting across development. *PLoS ONE*, *13*(6), e0198870. https://doi.org/10.1371/journal.pone.0198870

Perna, L. W., Ruby, A., Boruch, R. F., Wang, N., Scull, J., Ahmad, S., & Evans, C. (2014). Moving Through MOOCs: Understanding the Progression of Users in Massive Open Online Courses. *Educational Researcher*, *43*(9), 421–432. https://doi.org/10.3102/0013189X14562423

Phillips, N. E., Ralph, B. C. W., Carriere, J. S. A., & Smilek, D. (2016). Examining the influence of saliency of peer-induced distractions on direction of gaze and lecture recall. *Computers and Education*, *99*, 81–93. https://doi.org/10.1016/j.compedu.2016.04.006

Poerio, G. L., Totterdell, P., & Miles, E. (2013). Mind-wandering and negative mood: Does one thing really lead to another? *Consciousness and Cognition*, *22*(4), 1412–1421. https://doi.org/10.1016/j.concog.2013.09.012

Pomerantz, J. (2014). Data about the Metadata MOOC round 2. Retrieved July 5, 2015, from http://jeffrey.pomerantz.name/2014/10/data-about-the-metadata-mooc-round-2/

Pool, M. M., Koolstra, C. M., & Voort, T. H. A. (2003). The Impact of Background Radio and Television on High School Students' Homework Performance. *Journal of Communication*, *53*(1), 74–87. https://doi.org/10.1111/j.1460-2466.2003.tb03006.x

Qu, W., Ge, Y., Xiong, Y., Carciofo, R., Zhao, W., & Zhang, K. (2015). The relationship between mind wandering and dangerous driving behavior among Chinese drivers. *Safety Science*. Retrieved from http://www.sciencedirect.com/science/article/pii/S0925753515001083

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria; 2014. *URL Http://Www. R-Project. Org*. https://doi.org/10.1007/978-3-540-74686-7

Reeves, B., Lang, A., Kim, E. Y., & Tatar, D. (1999). The Effects of Screen Size and Message Content on Attention and Arousal. *Media Psychology*, *1*(1), 49–67. https://doi.org/10.1207/s1532785xmep0101_4

Reichle, E. D., Reineberg, A. E., & Schooler, J. W. (2010). Eye movements during mindless reading. *Psychological Science*, *21*(9), 1300–1310. https://doi.org/10.1177/0956797610378686

Reinten, J., Braat-Eggen, P. E., Hornikx, M., Kort, H. S. M., & Kohlrausch, A. (2017). The indoor sound environment and human task performance: A literature review on the role of room acoustics. *Building and Environment*, *123*, 315–332. https://doi.org/10.1016/j.buildenv.2017.07.005

Richardson, Jennifer C., P. D., & Swan, Karen, P. D. (2003). *Examing Social Presence in Online Courses in Relation to Students' Perceived Learning and Satisfaction*. Retrieved from https://www.ideals.illinois.edu/handle/2142/18713

Richardson, J. C., Maeda, Y., Lv, J., & Caskurlu, S. (2017). Social presence in relation to students' satisfaction and learning in the online environment: A meta-analysis. *Computers in Human Behavior*, *71*, 402–417. https://doi.org/10.1016/j.chb.2017.02.001

Rigby, J. M., Brumby, D. P., Cox, A. L., & Gould, S. J. J. (2016). Watching movies on netflix. *Proceedings of the 18th International Conference on Human-Computer Interaction with*

Mobile Devices and Services Adjunct - MobileHCI '16, 714–721. https://doi.org/10.1145/2957265.2961843

Risko, E. F., Anderson, N. C., 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. https://doi.org/10.1002/acp.1814

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: Mind wandering and computer use during lectures. *Computers & Education*, *68*, 275–283. https://doi.org/10.1016/j.compedu.2013.05.001

Risko, E. F., Laidlaw, K., Freeth, M., Foulsham, T., & Kingstone, A. (2012). Social attention with real versus reel stimuli: toward an empirical approach to concerns about ecological validity. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00143

Robison, M. K., & Unsworth, N. (2015). Working Memory Capacity Offers Resistance to Mind-Wandering and External Distraction in a Context-Specific Manner. *Applied Cognitive Psychology*, *29*(5), n/a-n/a. https://doi.org/10.1002/acp.3150

Rouder, J., Morey, R., Speckman, P., & Province, J. (2012). Default Bayes Factor for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/J.JMP.2012.08.001

Ruby, F. J. M., Smallwood, J., Engen, H., & Singer, T. (2013). How Self-Generated Thought Shapes Mood—The Relation between Mind-Wandering and Mood Depends on the Socio-Temporal Content of Thoughts. *PLoS ONE*, *8*(10), e77554. https://doi.org/10.1371/journal.pone.0077554

Sankaran, S. R., & Bui, T. (2001). Impact of Learning Strategies and Motivation on Performance: A Study in Web-Based Instruction. *Journal of Instructional Psychology*, *28*(3), 191. Retrieved from https://www.questia.com/library/journal/1G1-79370574/impact-of-learning-strategies-and-motivation-on-performance

Sayette, M. A., Reichle, E. D., & Schooler, J. W. (2009). Lost in the sauce: the effects of alcohol on mind wandering. *Psychological Science*, *20*(6), 747–752. https://doi.org/10.1111/j.1467-9280.2009.02351.x

Sayette, M. A., Schooler, J. W., & Reichle, E. D. (2010). Out for a smoke: the impact of cigarette craving on zoning out during reading. *Psychological Science*, *21*(1), 26–30. https://doi.org/10.1177/0956797609354059

Schacter, D. L., & Szpunar, K. K. (2015). Enhancing attention and memory during video-recorded lectures. *Scholarship of Teaching and Learning in Psychology*, *1*(1), 60–71. https://doi.org/10.1037/stl0000011

Schad, D. J., Nuthmann, A., & Engbert, R. (2012). Your mind wanders weakly, your mind wanders deeply: objective measures reveal mindless reading at different levels. *Cognition*, *125*(2), 179–194. https://doi.org/10.1016/j.cognition.2012.07.004

Schoen, J. R. (1970). Use of consciousness sampling to study teaching methods. *The Journal of Educational Research*, *63*(9), 387–390. https://doi.org/10.2307/27536016

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2004). Visual Metacognition in Adults and Children. In D. T. Levin (Ed.), *Thinking and Seeing: Visual metacognition in adults and children*.

Schooler, J. W., Smallwood, J., Christoff, K., Handy, T. C., Reichle, E. D., & Sayette, M. A. (2011). Meta-awareness, perceptual decoupling and the wandering mind. *Trends in Cognitive Sciences*, *15*(7), 319–326. https://doi.org/10.1016/j.tics.2011.05.006

Seli, P., Carriere, J. S. A. A., & Smilek, D. (2015). Not all mind wandering is created equal:

dissociating deliberate from spontaneous mind wandering. *Psychological Research*, *79*(5), 750–758. https://doi.org/10.1007/s00426-014-0617-x

Seli, P., Carriere, J. S. A., Levene, M., & Smilek, D. (2013). How few and far between? Examining the effects of probe rate on self-reported mind wandering. *Frontiers in Psychology*, *4*(July), 430. https://doi.org/10.3389/fpsyg.2013.00430

Seli, P., Carriere, J. S. A., Thomson, D. R., Cheyne, J. A., Martens, K. A. E., & Smilek, D. (2014). Restless Mind, Restless Body. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 660–668.

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015). Motivation, Intentionality, and Mind Wandering: Implications for Assessments of Task-Unrelated Thought. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, (MARCH). https://doi.org/10.1037/xlm0000116

Seli, P., Kane, M. J., Smallwood, J., Schacter, D. L., Maillet, D., Schooler, J. W., & Smilek, D. (2018). Mind-Wandering as a Natural Kind: A Family-Resemblances View. *Trends in Cognitive Sciences*, *22*(6), 479–490. https://doi.org/10.1016/J.TICS.2018.03.010

Seli, P., Konishi, M., Risko, E. F., & Smilek, D. (2018). The role of task difficulty in theoretical accounts of mind wandering. *Consciousness and Cognition*, *65*, 255–262. https://doi.org/10.1016/j.concog.2018.08.005

Seli, P., Ralph, B. C. W., Risko, E. F., W. Schooler, J., Schacter, D. L., & Smilek, D. (2017). Intentionality and meta-awareness of mind wandering: Are they one and the same, or distinct dimensions? *Psychonomic Bulletin & Review*, *24*(6), 1808–1818. https://doi.org/10.3758/s13423-017-1249-0

Seli, P., Risko, E. F., & Smilek, D. (2016). On the Necessity of Distinguishing Between

Unintentional and Intentional Mind Wandering. *Psychological Science*, *27*(5), 685–691. https://doi.org/10.1177/0956797616634068

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. https://doi.org/10.1016/j.tics.2016.05.010

Seli, P., Schacter, D. L., Risko, E. F., & Smilek, D. (2017). Increasing participant motivation reduces rates of intentional and unintentional mind wandering. *Psychological Research*. https://doi.org/10.1007/s00426-017-0914-2

Seli, P., Wammes, J. D., Risko, E. F., & Smilek, D. (2015). On the relation between motivation and retention in educational contexts: The role of intentional and unintentional mind wandering. *Psychonomic Bulletin & Review*, *23*(4), 1280–1287. https://doi.org/10.3758/s13423-015-0979-0

Shapiro, S. L., Carlson, L. E., Astin, J. A., & Freedman, B. (2006). Mechanisms of mindfulness. *Journal of Clinical Psychology*, *62*(3), 373–386. https://doi.org/10.1002/jclp.20237

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 Word Solution. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2160588

Sinnett, S., Spence, C., & Soto-Faraco, S. (2007). Visual dominance and attention: The Colavita effect revisited. *Perception and Psychophysics*, *69*(5), 673–686. https://doi.org/10.3758/BF03193770

Smallwood, J., Baracaia, S. F., Lowe, M., & Obonsawin, M. (2003). Task unrelated thought whilst encoding information. *Consciousness and Cognition*, *12*(3), 452–484. https://doi.org/10.1016/S1053-8100(03)00018-7

Smallwood, J., Fishman, D. J., & Schooler, J. W. (2007). Counting the cost of an absent mind:

Mind wandering as an underrecognized influence on educational performance. *Psychonomic Bulletin & Review*, *14*(2), 230–236. https://doi.org/10.3758/BF03194057

Smallwood, J., Fitzgerald, A., Miles, L. K., & Phillips, L. H. (2009). Shifting moods, wandering minds: Negative moods lead the mind to wander. *Emotion*, *9*(2), 271–276. https://doi.org/10.1037/a0014855

Smallwood, J., McSpadden, M., & Schooler, J. W. (2007). The lights are on but no one's home: meta-awareness and the decoupling of attention when the mind wanders. *Psychonomic Bulletin & Review*, *14*(3), 527–533. https://doi.org/10.3758/BF03194102

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: the curious incident of the wandering mind. *Memory & Cognition*, *36*(6), 1144–1150. https://doi.org/10.3758/MC.36.6.1144

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition*, *18*(1), 118–125. https://doi.org/10.1016/j.concog.2008.11.004

Smallwood, J., O'Connor, R. C., Sudbery, M. V., & Obonsawin, M. (2007). Mind-wandering and dysphoria. *Cognition & Emotion*, *21*(4), 816–842. https://doi.org/10.1080/02699930600911531

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin*, *132*(6), 946–958. https://doi.org/10.1037/0033-2909.132.6.946

Smallwood, J., & Schooler, J. W. (2015). The Science of Mind Wandering: Empirically Navigating the Stream of Consciousness. *Annual Review of Psychology*, *66*(1), 487–518. https://doi.org/10.1146/annurev-psych-010814-015331

Smallwood, J., Schooler, J. W., Turk, D. J., Cunningham, S. J., Burns, P., & Macrae, C. N.

(2011). Self-reflection and the temporal focus of the wandering mind. *Consciousness and Cognition*, *20*(4), 1120–1126. https://doi.org/10.1016/j.concog.2010.12.017

Spence, C., & Driver, J. (1997). Cross-Modal Links in Attention Between Audition, Vision, and Touch: Implications for Interface Design. *International Journal of Cognitive Ergonomics.*, *1*(4), 351–373. Retrieved from http://psycnet.apa.org/record/1999-10777-005

Spence, C., Nicholls, M. E. R., & Driver, J. (2001). The cost of expecting events in the wrong sensory modality. *Perception & Psychophysics*, *63*(2), 330–336. https://doi.org/10.3758/BF03194473

Stawarczyk, D., Majerus, S., Maj, M., Van der Linden, M., & D'Argembeau, A. (2011). Mind-wandering: Phenomenology and function as assessed with a novel experience sampling method. *Acta Psychologica*, *136*(3), 370–381. https://doi.org/10.1016/j.actpsy.2011.01.002

Stawarczyk, D., Majerus, S., Maquet, P., & D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. *PloS One*, *6*(2), e16997. https://doi.org/10.1371/journal.pone.0016997

Struk, A. A., Carriere, J. S. A., Cheyne, J. A., & Danckert, J. (2017). A Short Boredom Proneness Scale. *Assessment*, *24*(3), 346–359. https://doi.org/10.1177/1073191115609996

Sung, E., & Mayer, R. E. (2012). Five facets of social presence in online distance education. *Computers in Human Behavior*, *28*(5), 1738–1747. https://doi.org/10.1016/j.chb.2012.04.014

Szpunar, K. K., Khan, N. Y., & Schacter, D. L. (2013). Interpolated memory tests reduce mind wandering and improve learning of online lectures. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6313–6317.

https://doi.org/10.1073/pnas.1221764110

Szpunar, K. K., Moulton, S. T., & Schacter, D. L. (2013). Mind wandering and education: from the classroom to online learning. *Frontiers in Psychology*, *4*(August), 495. https://doi.org/10.3389/fpsyg.2013.00495

Takatalo, J., Häkkinen, J., Kaistinen, J., & Nyman, G. (2010). Presence, involvement, and flow in digital games. In *Evaluating user experience in games* (pp. 23–46).

Theeuwes, J., & Van der Stigchel, S. (2006). Faces capture attention: Evidence from inhibition of return. *Visual Cognition*, *13*(6), 657–665. https://doi.org/10.1080/13506280500410949

Thomson, D. R., Seli, P., Besner, D., & Smilek, D. (2014). On the link between mind wandering and task performance over time. *Consciousness and Cognition*, *27*, 14–26. https://doi.org/10.1016/j.concog.2014.04.001

Tunnell, G. B. (1977). Three dimensions of naturalness: An expanded definition of field research. *Psychological Bulletin*, *84*(3), 426–437. https://doi.org/10.1037/0033-2909.84.3.426

Unsworth, N., Brewer, G. A., & Spillers, G. J. (2012). Variation in cognitive failures: An individual differences investigation of everyday attention and memory failures. *Journal of Memory and Language*, *67*(1), 1–16. https://doi.org/10.1016/j.jml.2011.12.005

Unsworth, N., & McMillan, B. D. (2013). Mind wandering and reading comprehension: examining the roles of working memory capacity, interest, motivation, and topic experience. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(3), 832–842. https://doi.org/10.1037/a0029669

Unsworth, N., & McMillan, B. D. (2014). Similarities and differences between mind-wandering and external distraction: A latent variable analysis of lapses of attention and their relation to cognitive abilities. *Acta Psychologica*, *150*, 14–25. https://doi.org/10.1016/j.actpsy.2014.04.001

Unsworth, N., & McMillan, B. D. (2017). Attentional disengagements in educational contexts: a diary investigation of everyday mind-wandering and distraction. *Cognitive Research: Principles and Implications*, *2*(1), 32. https://doi.org/10.1186/s41235-017-0070-7

Unsworth, N., McMillan, B. D., Brewer, G. a, & Spillers, G. J. (2012). Everyday attention failures: an individual differences investigation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(6), 1765–1772. https://doi.org/10.1037/a0028075

Unsworth, N., & Robison, M. K. (2016). The influence of lapses of attention on working memory capacity. *Memory & Cognition*, *44*(2), 188–196. https://doi.org/10.3758/s13421-015-0560-0

van Vugt, M. K., & Broers, N. (2016). Self-Reported Stickiness of Mind-Wandering Affects Task Performance. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00732

Varao-Sousa, T.L., & Kingstone, A. (2015). Memory for lectures: How lecture format impacts the learning experience. *PLoS ONE*, *10*(11). https://doi.org/10.1371/journal.pone.0141587

Varao-Sousa, T.L., Solman, G. J. F., & Kingstone, A. (2017). Re-Reading after Mind Wandering. *Canadian Journal of Experimental Psychology*, *71*(3). https://doi.org/10.1037/cep0000123.supp

Varao-Sousa, Trish L., & Kingstone, A. (2018). Are mind wandering rates an artifact of the probe-caught method? Using self-caught mind wandering in the classroom to test, and reject, this possibility. *Behavior Research Methods*, 1–8. https://doi.org/10.3758/s13428-018-1073-0

Varao Sousa, T. L., Carriere, J. S. A., & Smilek, D. (2013). The way we encounter reading material influences how frequently we mind wander. *Frontiers in Psychology*, *4*, 892. https://doi.org/10.3389/fpsyg.2013.00892

Video Game Industry - Statistics & Facts | Statista. (n.d.). Retrieved October 18, 2018, from https://www.statista.com/topics/868/video-games/

Wammes, J. D., Boucher, P. O., Seli, P., Cheyne, J. A., & Smilek, D. (2016). Mind wandering during lectures I: Changes in rates across an entire semester. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 13–32. https://doi.org/10.1037/stl0000053

Wammes, J. D., Seli, P., Cheyne, J. A., Boucher, P. O., & Smilek, D. (2016). Mind wandering during lectures II: Relation to academic performance. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 33–48. https://doi.org/10.1037/stl0000055

Wammes, J. D., & Smilek, D. (2017). Examining the Influence of Lecture Format on Degree of Mind Wandering. *Journal of Applied Research in Memory and Cognition*, *6*(2), 174–184. https://doi.org/10.1016/j.jarmac.2017.01.015

Weinstein, Y. (2017). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behaviour Research Methods*, (December). https://doi.org/10.3758/s13428-017-0891-9

Weinstein, Y., De Lima, H. J., & van der Zee, T. (2018). Are you mind-wandering, or is your mind on task? The effect of probe framing on mind-wandering reports. *Psychonomic Bulletin & Review*, *25*(2), 754–760. https://doi.org/10.3758/s13423-017-1322-8

Wilson, K. E., Martinez, M., Mills, C., D’Mello, S., Smilek, D., & Risko, E. F. (2018). Instructor presence effect: Liking does not always lead to learning. *Computers & Education*, *122*, 205–220. https://doi.org/10.1016/J.COMPEDU.2018.03.011

Wilson, K., & Korn, J. H. (2007). Attention During Lectures: Beyond Ten Minutes. *Teaching of*

Psychology, *34*(2), 85–89. https://doi.org/10.1080/00986280701291291

Wuensch, K. L., Aziz, S., Ozan, E., Kishore, M., & Tabrizi, M. H. N. (2008). Pedagogical Characteristics of Online and Face-to-Face Classes. *World Conference on E-Learning in Corporate, Government, Healthcare, and Heather Education*, *7*, 523–532.

Xu, D., & Jaggars, S. S. (2013). The impact of online learning on students' course outcomes: Evidence from a large community and technical college system. *Economics of Education Review*, *37*, 46–57. https://doi.org/10.1016/j.econedurev.2013.08.001

Yanko, M. R., & Spalek, T. M. (2013). Route familiarity breeds inattention: a driving simulator study. *Accident Analysis & Prevention*, *57*, 80–86. https://doi.org/10.1016/j.aap.2013.04.003

Yildirim, Z., & Kilis, S. (2019). Posting patterns of students' social presence, cognitive presence, and teaching presence in online learning. *Online Learning Journal*, *23*(2), 179–195. https://doi.org/10.24059/olj.v23i2.1460

Young, M. S., Robinson, S., & Alberts, P. (2009). Students pay attention! *Active Learning in Higher Education*, *10*(1), 41–55. https://doi.org/10.1177/1469787408100194

Zhang, H., Lin, L., Zhan, Y., & Ren, Y. (2016). The Impact of Teaching Presence on Online Engagement Behaviors. *Journal of Educational Computing Research*, *54*(7), 887–900. https://doi.org/10.1177/0735633116648171