

**Capturing Dynamic Fluctuations in Remembered Momentary Attentional States**

by

Samantha Ayers-Glassey

A thesis

presented to the University of Waterloo

in fulfilment of the

thesis requirement for the degree of

Master of Arts

in

Psychology

Waterloo, Ontario, Canada, 2022

© Samantha Ayers-Glassey 2022

### **Author's Declaration**

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

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

## Statement of Contributions

**Samantha Ayers-Glassey** was the primary author of the manuscript, conceived the experimental design, and analyzed the data.

**Effie J. Pereira** programmed the experiment and assisted with conceiving the experimental design, analyzing the data, and writing the manuscript.

**Jeffrey D. Wammes** assisted with conceiving the experimental design.

**Daniel Smilek** guided the research process including the experimental design, data analysis, and writing of the manuscript.

## Abstract

States of attention such as engagement and mind-wandering fluctuate dynamically over time. Although aggregate retrospective reports of attention have been used in prior research, the full nature of people's *remembered momentary attentional states* (RMASs) – i.e., how attentive they were at a specific moment in time – remains largely unexplored. Across two experiments, participants ( $N_{E1} = 102$ ,  $N_{E2} = 97$ ) watched online video lectures while being intermittently probed to report their subjective levels of attentional engagement (E1) or mind-wandering (E2). After watching the lecture, participants were then presented with short excerpts from it and were asked to recall their level of attentional engagement or mind-wandering when they initially watched the excerpt within the lecture. When reporting on attentional engagement (E1), cross-correlation analyses revealed concordance between the temporal dynamics of participants' in-the-moment and retrospective ratings. However, when reporting on mind-wandering (E2), there was only concordance for videos that were considered less engaging. Contrasts across experiments revealed that the overlap between in-the-moment and retrospective ratings were significantly lower when participants reported mind-wandering compared to attentional engagement. Together, our findings suggest that people may have enduring RMASs for general attentional engagement but less-robust memories for the specific experience of mind-wandering.

## **Acknowledgments**

I would like to thank my supervisor, Dr. Daniel Smilek, for his guidance and enthusiasm throughout this project; my collaborators, Dr. Effie J. Pereira and Dr. Jeffrey D. Wammes for their contributions to the design, programming, and analysis of these studies; my labmates – especially Lydia and Alyssa – for their support with transitioning into a new city during lockdown and with navigating the demands of a new degree online; and my readers, Dr. Michael Dixon and Dr. Evan Risko, for their time and feedback on this thesis.

Additionally, I thank my parents for their ongoing unconditional support in all my pursuits, and for being incredible role models both within and outside of academia; Patt for her advice and support through transitioning into graduate school and adulthood; and Tom for his unwavering encouragement and reminders to eat lunch when I was too busy analyzing data or writing up results to remember.

This research was conducted at the University of Waterloo and was supported by the Discovery grant from the National Sciences and Engineering Research Council of Canada (NSERC) awarded to Dr. Daniel Smilek. It was also supported by the Canada Graduate Scholarship for the Master's program (CGS-M), which I received from NSERC.

## **Dedication**

To the Barbaras with love, gratitude, and admiration.

## Table of Contents

Author's Declaration.....	ii
Statement of Contributions .....	iii
Abstract .....	iv
Acknowledgments.....	v
Dedication.....	vi
List of Figures .....	viii
List of Tables .....	ix
Introduction.....	1
Experiment 1: Attentional Engagement.....	13
Method .....	17
Participants .....	17
Materials & Stimuli .....	17
Procedure .....	19
Results .....	20
Discussion .....	27
Experiment 2: Mind-Wandering .....	28
Method .....	29
Participants .....	29
Materials & Stimuli .....	30
Procedure .....	30
Results .....	31
Discussion .....	37
Comparing Attentional Engagement and Mind-Wandering .....	39
General Discussion .....	41
References.....	48
Appendices.....	57

## List of Figures

**Figure 1.** A sample participant's less-engaging video condition results depicting the cross-correlation analytic procedure used.

**Figure 2.** A sample participant's cross-correlations for the less-engaging video condition.

**Figure 3.** The video-viewing task and video-stimulated recall task procedures for Experiment 1.

**Figure 4.** Mean engagement probe responses for the 'less-engaging' Yale lecture and 'more-engaging' TEDx Talk video conditions as a function of immediate and retrospective probe types.

**Figure 5.** Mean cross-correlation scores as a function of time lag from the (a) less-engaging and (b) more-engaging video conditions of Experiment 1.

**Figure 6.** Mean immediate and retrospective (a) intentional and (b) unintentional mind-wandering probe responses for the less-engaging and more-engaging video conditions of Experiment 2.

**Figure 7.** Mean cross-correlation scores as a function of time lag from the (a) less-engaging intentional mind-wandering, (b) less-engaging unintentional mind-wandering, (c) more-engaging intentional mind-wandering, and (d) more-engaging unintentional mind-wandering conditions of Experiment 2.

**Figure 8.** Mean cross-correlation scores as a function of time lag from the standardized (a) Experiment 1 less-engaging, (b) Experiment 2 less-engaging, (c) Experiment 1 more-engaging, and (d) Experiment 2 more-engaging conditions.

## List of Tables

**Table 1.** Correlations between attentional-engagement probe responses and content-recall quiz results for Experiment 1.

**Table 2.** Correlations between mind-wandering probe responses and content-recall quiz results for Experiment 2.

## Introduction

Attention can vary between states of engagement with the external environment and states of engagement with internal thoughts (i.e., mind wandering), dynamically fluctuating between external and internal content on a moment-to-moment basis. Suppose you were a student attending a 90-minute university lecture. Naturally, your attention would ebb and flow throughout the duration of the lecture, perhaps being completely engrossed in the presentation one moment but distracted by unrelated thoughts (i.e., mind-wandering) in another; these fluctuations over time are referred to as *temporal dynamics*. Prior studies have demonstrated that attentional states can vary during all activities, including during periods of rest (Christoff et al., 2016; Delorme & Brandmeyer, 2019; Hasenkamp et al., 2012; Tusche et al., 2014), during leisure activities (Feng et al., 2013; Killingsworth & Gilbert, 2010; Kopp et al., 2016), and during work activities (Risko et al., 2013; Szpunar, Khan, et al., 2013; Szpunar, Moulton, et al., 2013). Optimal performance on everyday tasks requires levels of attentional engagement to match task demands, and if attentional engagement is suboptimal, performance errors can occur (Carriere et al., 2008; Ralph et al., 2017; Randall et al., 2014; Thomson et al., 2015).

While researchers have begun exploring how attentional engagement unfolds over time (e.g., Brosowsky et al., 2020; Thomson et al., 2014; Wammes et al., 2016), very little is known about people's memories for the temporal dynamics of their prior attentional states. We can imagine that as a student that attends lectures, you might later be studying the lecture's content and recall that your attention was not completely focused during a specific point in the presentation. Such a memory might be useful because you might then decide to allocate more study time to the specific lecture content to which you failed to attend. The current thesis

explores the possibility that individuals might have enduring memories for momentarily fluctuating attentional states, specifically states of attentional engagement and mind-wandering.

### **Capturing Self-Reported Attentional Experiences**

Several methods exist for capturing dynamic fluctuations in self-reported attentional states. These include *immediate thought probes*, wherein participants are directly asked about their attentional states intermittently throughout a task (Christoff et al., 2009; Risko et al., 2012; Wammes, Seli, et al., 2016); *aggregate retrospective questionnaires*, wherein participants are asked to report the number of attentional experiences that occurred across an entire task after they have completed it (Barron et al., 2011; Kane et al., 2021; Smallwood et al., 2004); and *short-term retrospective thought probes*, wherein participants are asked to report whether an attentional experience had occurred since the preceding probe (Antrobus, 1968; Giambra, 1995; Shaw & Giambra, 1993). Each of these methods captures a unique aspect of the temporal dynamics of attention with varying levels of delay between the initial attentional experience and when individuals report on that experience. Thus, in studies using these methods there is implied examination of memory for attentional states. In what follows, each of these three methods will be considered in greater detail.

#### ***Immediate Probe Reports***

The goal of immediate thought probing is to tap into an individual's ongoing, in-the-moment experiences with as little demand on memory as possible by asking about some aspect of their attentional state experienced right before the presentation of a thought probe. Although the wording of probes can vary between studies, the general procedure remains the same with probes being presented several times throughout a task. For example, probes in a study by Smallwood et al. (2008) asked participants, "*Just prior to being asked, was your attention on- or*

*off-task?*” with response options of ‘tuning out’, ‘zoning out’, or ‘on task’. Probes in a study by Christoff et al. (2009) asked participants, “*Where was your attention focused just before the probe?*” with response options from 1 (on-task) to 7 (off-task) and was followed up with, “*How aware were you of where your attention was focused?*” with response options from 1 (aware) to 7 (not aware). As another example, probes in a study by Stawarczyk et al. (2011) asked participants to characterize the ongoing conscious experience they had just prior to the probe with response options including ‘on-task’, ‘task-related interference’ (i.e., interfering thoughts related to the appraisal of the task), ‘external distractions’, or ‘stimulus-dependent and task-unrelated thoughts’ (i.e., mind-wandering or daydreams). Using a variety of probes such as those mentioned above, researchers have been able answer many different questions regarding individuals’ attentional experiences.

Importantly, immediate thought probes allow researchers to collect time-sensitive data that can be used to deduce the temporal dynamics of an individual’s attention, i.e., how one’s attention fluctuates over time (Thomson et al., 2014). Specifically, researchers have observed ‘time-on-task effects’ using immediate probing. For example, mind-wandering (probed as, “*Were you mind-wandering?*” with response options of ‘yes’ or ‘no’) has been found to increase as a function of time from the first half to the second half of a video-recorded university lecture (Risko et al., 2012). Furthermore, changes in mind-wandering over time on task (probed as asking whether participants were ‘on-task’ or ‘off-task’) have been found to significantly predict changes in visual-search accuracy (Thomson et al., 2014). Therefore, employing immediate probes throughout a task provides researchers with many different time-specific data points that can be analyzed to unveil the temporal dynamics of an individual’s self-reported fluctuating attention across a task.

Immediate probing is the most commonly employed method for capturing self-reports in studies examining the temporal dynamics of attention due to several of its benefits. First, immediate probes can be employed across a diverse set of activities. Studies have examined the temporal dynamics of attention throughout trial-based attentional tasks such as the sustained attention to response task (SART; Christoff et al., 2009; Seli, Risko, & Smilek, 2016), flanker tasks (Kane et al., 2016; Kane et al., 2021), and the metronome-response task (MRT; Anderson et al., 2020; Seli, Cheyne, et al., 2013). Additionally, researchers have used immediate probing throughout free-flowing, unconstrained activities such as reading (Phillips et al., 2016; Soemer & Schiefele, 2019) and watching a video (Risko et al., 2012; Wammes & Smilek, 2017). Immediate probes have also been employed in naturalistic settings such as during live classroom lectures (Wammes, Boucher, et al., 2016; Wammes et al., 2019) and throughout everyday activities (Kane et al., 2007; McVay et al., 2009). Therefore, this method allows for the examination of dynamically fluctuating attentional states across many different activities.

Second, immediate probes require little memory for an individual's attentional state as they ask about one's immediately preceding experience (Myin-Germeys et al., 2009; Robison et al., 2019; Weinstein, 2018). As such, immediate probe responses can typically predict task performance. For example, mind-wandering probe reports have been found to correlate negatively with target accuracy on the SART (McVay & Kane, 2009; 2012), behavioural variability on the MRT (Seli, Cheyne, et al., 2013), recall of lecture material (Risko et al., 2012), and text comprehension (Smallwood et al., 2008). Furthermore, immediate mind-wandering reports have been associated with pupil dilation (Unsworth & Robison, 2016) and specific brain activity (i.e., EEG; Baird et al., 2014). Thus, it can be assumed that individuals' immediate

reports of their ongoing attentional experiences are not meaningfully impacted by memory limitations (Weinstein, 2018).

Finally, immediate probes can be analyzed across different timeframes, such as individual points along a short timeseries or aggregated for broader examination. For example, Wammes, Boucher, et al. (2016) had undergraduate students respond to attention probes up to three times in each of 36 live lectures presented throughout a 12-week psychology course. This allowed for examinations of students' dynamic self-reported attention (a) over the course of an average lecture, (b) over the course of an average week, and (c) over the entire term. This demonstrates how the immediate-probing method can allow for precise comparisons of self-reported attentional states across various timeframes.

Overall, the immediate-probing method employs the shortest possible delay between an individual's attentional experience and when they report on that experience. Although there is inherently some form of retrospection involved in reporting on an experience occurring just prior to a probe, it is often assumed that since the delay is so short memory would have a minimal impact on the attentional state being probed.

### ***Aggregate Retrospective Reports***

Alternatively, aggregate retrospective reporting employs the longest delay of the three methods between an individual's attentional experience and their reported recall of that experience. This method involves individuals completing a task and then reporting on their *overall* levels of attention experienced throughout that task. For example, items from the Thinking Content subscale of the Dundee Stress State Questionnaire (DSSQ; Matthews et al., 1999; Matthews et al., 2013) are commonly used as a retrospective index of concentration and mind-wandering (e.g., Banks et al., 2015; Barron et al., 2011; Kane et al., 2021; Smallwood et

al., 2006). These items ask participants about the aggregated frequency that they remember thinking about different topics, both on-task and off-task, during a task on a scale from 1 (never) to 5 (very often). For example, “*I thought about personal worries*” and “*I thought about something that happened in the distant past*”. When assessing retrospective reports of attention, the DSSQ has been found to correlate positively with immediate mind-wandering probes responded to throughout a SART task ( $r = .36$ ; Kane et al., 2021), supporting some overlap between individuals’ in-the-moment and aggregated retrospective self-reports of attentional experiences.

One benefit of aggregate retrospective reports is that they can be employed at the end of a study without the need to interrupt an ongoing task, an issue that has been shown to impact immediate reporting of attentional experiences (Myin-Germeys et al., 2009; Seli, Carriere, et al., 2013). Seli, Carriere, et al. (2013) presented immediate thought probes throughout an attention task at varying intervals (between 5-25 probes total) and found a significant relation between the time between probes and proportions of participants’ probe-caught mind-wandering. Specifically, more frequently presented probes were associated with fewer mind-wandering reports, suggesting that probing participants too frequently on their attention in the moment might lead to altered self-reported mind-wandering either due to a lack of time to experience mind-wandering between probes or, more likely, reducing the likelihood that individuals will report on their mind-wandering experiences. Employing retrospective reports following task completion might eliminate the disruptiveness of immediate probing (Smallwood & Schooler, 2015).

While the aggregate retrospective-report method has some advantages, researchers have expressed concerns regarding the recall accuracy of retrospective reports. Kane et al. (2021) summarized the disadvantages of aggregated retrospective questionnaires, stating:

These questionnaires allow efficient data collection, but their validity as measures of individual differences in mind-wandering propensity rests on people's ability to notice their fleeting conscious experiences as they occur, to faithfully recall and aggregate them over long (typically unspecified) timescales, and then to accurately translate that aggregation into a relative frequency or agreement rating. (p. 2401)

The effect of having a longer delay between an immediate and aggregate retrospective report of attention, as well as the effect of having to report on the aggregate experience, has not been thoroughly examined. However, limitations with the aggregate retrospective report technique have been found in the context of other conscious states, such as affect. For example, Ellison et al. (2020) examined whether individuals could recall moment-to-moment self-esteem and stress states similarly to how they had originally rated those experiences in the moment. The researchers had participants complete 32 pairs of probes via smartphone every 15 minutes for eight hours during two days of regular routine: a self-esteem probe (*"Right now, I feel good about myself"*) and a stress probe (*"Right now, I feel stressed"*), with response options for both probes ranging from 0 (not at all) to 100 (extremely). Then, they had participants return to the lab approximately one week later to recall their average self-esteem and stress experiences on the same 0-100 scale. The authors found that participants recalled similar aggregate levels of self-esteem but higher aggregate levels of stress than they had originally reported. These results suggest that individuals' aggregated recall of some prior states (such as stress) can be inflated and therefore are not always representative of their immediately reported experiences.

Furthermore, retrospective reporting of affective and physiological experiences has been found to rely on saliency and recency effects. The *peak-end rule* describes how individuals typically report retrospectively on aggregate affective and physiological experiences based on the peak intensity (i.e., the most salient experience) and the end intensity (i.e., the most recent experience) rather than on an equally weighted average of the entire experience (Redelmeier & Kahneman, 1996). This has been shown to persist regardless of the length of an experience, a phenomenon referred to as *duration neglect* (see Alaybek et al., 2022 for a review and meta-analysis). For example, Redelmeier and Kahneman (1996) compared colonoscopy and lithotripsy patients' in-the-moment pain ratings (reported every 60 seconds via an electronic device as being between 0 [no pain] and 10 [extreme pain]) and their aggregate retrospective pain ratings (reported one hour after the clinical procedure as their 'total amount of pain experienced' between 0 [no discomfort] and 10 [awful discomfort]). They found strong correlations between the participants' aggregate retrospective pain ratings and the peak ( $r = .61$ ) and end pain ( $r = .44$ ), but no significant correlation with procedure duration ( $r = .12$ ). Furthermore, a simple multivariate model with peak pain and end pain as predictors of patients' aggregate retrospective pain reports was not significantly improved by adding total pain, average pain, initial pain, or duration as predictors. Thus, Redelmeier and Kahneman (1996) concluded that participants' retrospective reporting of the overall pain experienced during the two types of medical procedures were characterized by peak and end experiences, a finding that has been supported consistently in the literature (Alaybek et al., 2022). Therefore, reports at the end of a study likely impose a memory demand and a bias towards saliency and recency that is not apparent in immediate probing, and thus might be a less-accurate method of collecting self-reported attention than immediate probes.

### ***Short-Term Retrospective Probe Reports***

Short-term retrospective probing is a much less-common approach to capturing self-reported attentional experiences that combines the minimal delay of immediate probing and the recall of retrospective questionnaires while avoiding some of the weaknesses associated with retrospective reporting. This method involves participants responding to probes throughout a task, reporting whether they experienced mind-wandering in the interval between the last probe and the current probe. For example, Antrobus (1968) probed participants every 15 seconds, asking whether they remembered mind-wandering at least once since the previous probe (with response options of ‘yes’ or ‘no’); Shaw and Giambra (1993) probed participants every 25 seconds, asking participants to report if they had mind-wandered since the previous probe (with responses reported via keys denoted ‘spontaneous’ [ins key] or ‘deliberate’ [+ key]); and Giambra (1995) had participants respond to probes every 29 seconds, asking participants to report whether they had mind-wandered since the last probe (with responses reported via a key denoted ‘TUT’ [task-unrelated thought; i.e., mind-wandering]). Although these might seem like very short intervals, they still require participants to report on their recall of attentional experiences since the last time they responded.

### **Remembered Momentary Attentional States**

Although the aforementioned methods have been used in prior research, the full nature of people’s memories for how attentive they were at a specific moment in time – here referred to as *remembered momentary attentional states* (RMASs) – remains largely unexplored. Specifically, studies have yet to examine people’s ability to recall their moment-to-moment attentional fluctuations after a task has been completed. Considering individuals typically have some control over their attentional experiences (e.g., the capability to intentionally mind-wander; Seli, Risko,

& Smilek, 2016; Seli, Risko, Smilek, et al., 2016), recall of specific past attentional experiences could be useful for informing how one might allocate one's attention when a similar situation is encountered in the future. For example, an individual might recall that they typically have a harder time paying attention near the end of a lecture and so plan to exert more effort to sustain their attention later in a lecture. Thus, there is reason to believe that people may have durable memories of specific prior attentional states, and understanding the nature of these memories could shed light on how people could use such knowledge to guide their attention in the future.

Assessing RMASSs requires approaching the retrospective measurement of attentional experiences differently than prior retrospective methods have. This is due to the need to capture the precise temporal dynamics of attention after a task has been completed, which cannot be met by aggregate estimates of attentional engagement. One possible approach is *stimulated recall*, a method wherein participants' retrospective experiences are measured with the aid of audio or video stimuli from the original event (Bloom, 1953; Calderhead, 1981; Gass & Mackey, 2016; Lyle, 2003). For example, communication researchers have used this method to examine second-language students' dynamic self-ratings of anxiety throughout an oral presentation to their class by video-recording the presentations and having participants watch the recording as a recall cue while rating their initial experiences of anxiety (Gregersen et al., 2014). These dynamic ratings of anxiety were found to have strong convergence with heart-rate data that was collected during the recorded presentations, demonstrating that participants' cued retrospective reports of anxiety were consistent with a physiological measure of anxiety. Stimulated recall has also been used to capture memories of students' thoughts during classroom lectures (O'Brien, 1993), children's perceptions of their learning (Morgan, 2007), teachers' pedagogical beliefs (Meade & McMeniman, 1992), musicians' techniques for successful collaboration (Dempsey, 2010),

physicians' memory for clinical encounters (see Sinnott et al., 2017 for a review), and individuals' working-memory constraints due to cognitive load (Beers et al., 2006). Given the success of this method, it seems reasonable to posit that stimulated recall could be used to cue individuals' memory of prior experiences for the purpose of reporting on their specific RMASSs.

Stimulated recall has several strengths when assessing subjective experiences retrospectively (Bloom, 1953; Gass & Mackey, 2016; Lyle, 2003). First, like immediate probing, stimulated recall can be used quite widely across a number of activities given that there is no direct interruption to the initial behaviour itself. Second, also like immediate probing, stimulated recall can be utilized to recall as many time points as necessary allowing for specific probing of dynamic and free-flowing experiences. Finally, unlike aggregate retrospective reporting, stimulated recall does not require participants to mentally average specific experiences to arrive at a representative overall experience that may be influenced by forgetting (Ellison et al., 2020) or particularly salient but non-representative moments (Alaybek et al., 2022; Redelmeier & Kahneman, 1996). Instead, stimulated recall involves a re-presentation of a particular moment from the original task to cue an individual's recall of their experiences during that specific timepoint. For these complementary reasons, this method could prove particularly beneficial for assessing RMASSs with a high degree of temporal specificity across a number of diverse activities, especially considering that stimulated recall could be effectively matched with immediate subjective reports to make immediate and retrospective probes more directly comparable.

## **The Present Studies**

Across two experiments, we examined the nature of individuals' retrospective reporting of dynamic fluctuations in their momentary attentional states, focusing on the broader experience

of attentional engagement and the more specific experience of mind-wandering. To do so, we utilized immediate thought probes during a video-viewing task to capture participants' in-the-moment assessments of attention (immediate reports). Following this, we implemented video-stimulated recall by presenting participants with short clips from the video they had just viewed and asking them to estimate their attention when they first viewed that section of the video (retrospective reports). Immediate and retrospective reports were then assessed for temporal overlap using cross-correlation analysis to quantitatively evaluate the correspondence between the two timeseries. Experiment 1 assessed the congruence between participants' immediate and retrospective reports of attentional engagement. Experiment 2 then assessed the congruence between participants' immediate and retrospective reports of mind-wandering. The results of the two RMASs were then compared to assess whether there was greater correspondence between immediate and retrospective probe reports when assessing attentional engagement than when assessing mind-wandering.

## Experiment 1: Attentional Engagement

In Experiment 1, we examined the overarching RMAS of *attentional engagement* by assessing the degree of correspondence between individuals' immediate and retrospective probe reports of this attentional state. When one is in a state of engagement, attention is primarily focused on the task at hand (Weinstein et al., 2018). We chose to examine memory for attentional engagement first because we assumed that it is a broad attentional experience and so could be easier to recall than a more specific state such as mind-wandering. To examine the nature of individuals' memory for their initial attentional engagement while viewing a video lecture, we compared individuals' immediate reports of attentional engagement captured during the initial viewings of the lecture to their retrospective recall cued by re-presentations of segments of the lecture. Two video lectures were selected to assess whether similar outcomes would be observed in lectures that were inherently less or more engaging: a Yale lecture (less engaging) and a TEDx Talk (more engaging). Additionally, we had participants complete a short quiz following the tasks to assess their recall of the video content. Finally, we assessed the concordance between the temporal dynamics of participants' immediate and retrospective reports of attentional engagement for both video conditions.

Our analyses were aimed at addressing several issues. First, we assessed the average immediate and average retrospective reports as a function of video engagingness to evaluate whether there is any overall bias in recall of attentional engagement relative to immediate reports and to confirm that the manipulation of video engagingness was successful. Second, we assessed the relations between immediate reports of attentional engagement, retrospective reports of attentional engagement, and recall of the lecture content; robust correlations among these measures would be consistent with the notion that they each index momentary levels of

attentional engagement. Finally, and most importantly, we examined the degree of correspondence in the temporal dynamics of the immediate and retrospective measures. In this analysis we also examined whether the degree of correspondence between immediate and retrospective measures depended on the engagingness of the video viewed.

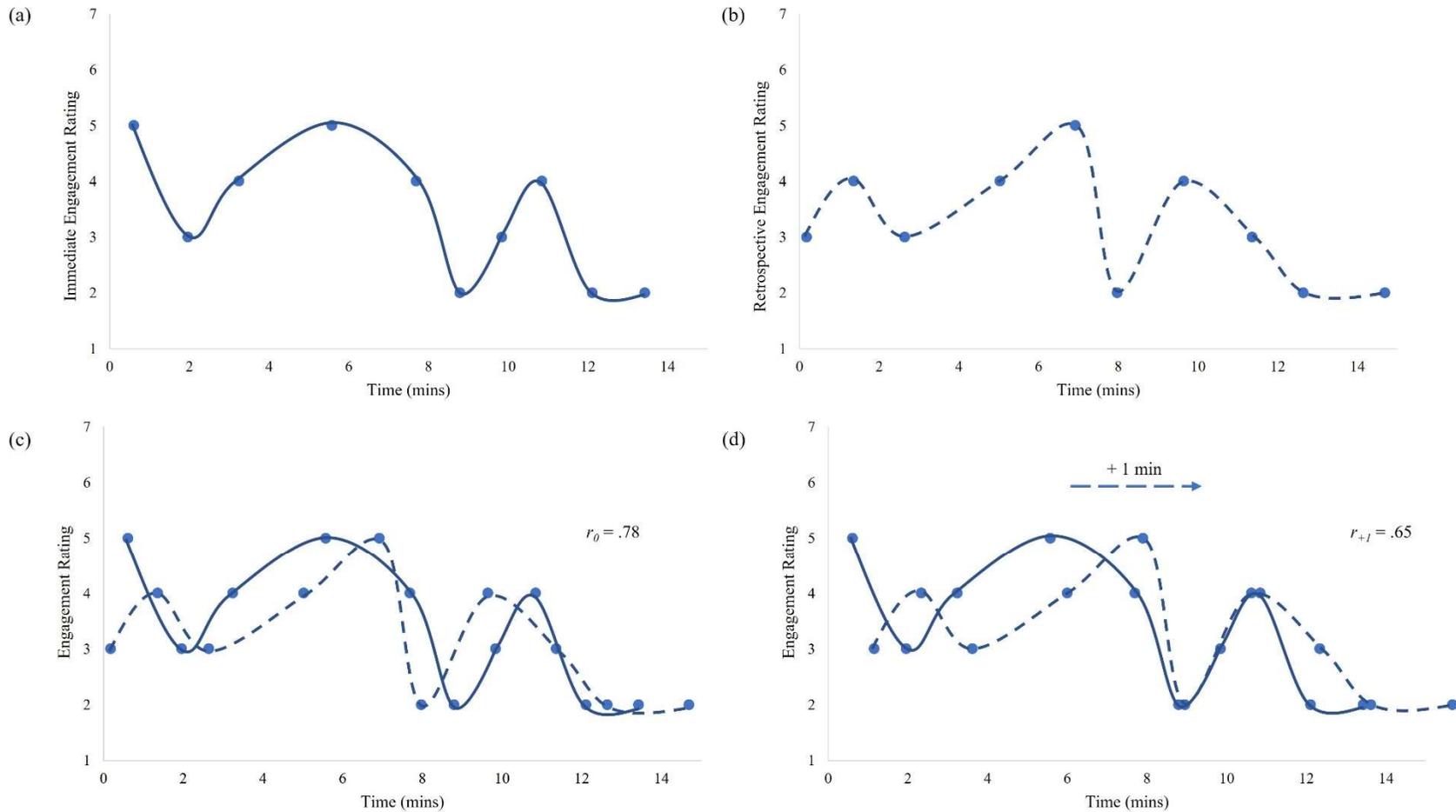
To compare the temporal dynamics of the immediate and retrospective measures, we used *cross-correlation*, an analytic method that calculates the degree of overlap between the timeseries that underlie the two measures (Brunsdon & Skinner, 1987; Derrick & Thomas, 2004; Gregson, 2014; Heath, 2000). We used this technique because the time stamps for the immediate and retrospective thought probes never overlapped based on video runtime, which precluded a direct comparison of the probes across measures. The cross-correlation analysis was applied to each video condition separately.

As a concrete example of a single cross-correlation analysis, consider the sample data depicted in Figure 1. Here, we can assume that discrete responses to thought probes of attentional engagement are merely time points along a continuous underlying timeseries of attentional engagement. In this manner, we can consider immediate probe responses to be data points along an *immediate timeseries* (Figure 1a), which represents the reference set or ground truth of an individual's attentional engagement over time. Similarly, retrospective probe responses can be considered data points along a *retrospective timeseries* (Figure 1b), which represents the evaluation set that is meant to be compared against the reference set.

To examine whether immediate and retrospective timeseries are representative of the same underlying fluctuations in attentional engagement, cross-correlation assesses the precise temporal relationship between the two by calculating whether they occur in a time-locked pattern relative to one another. For this, we first assess the cross-correlation score between the

**Figure 1.**

*A Sample Participant's Less-Engaging Video Condition Results Depicting the Cross-Correlation Analytic Procedure Used.*

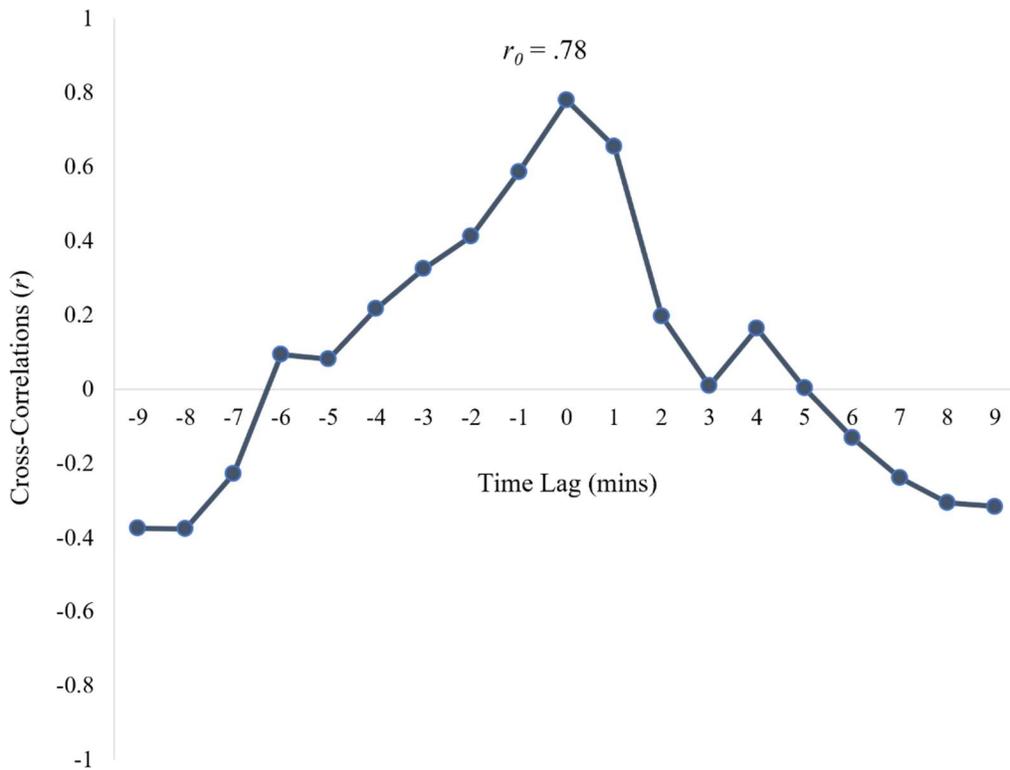


*Note:* Since (a) immediate probe responses and (b) retrospective probe responses are assumed to represent discrete points along a continuum of attentional engagement, we converted these responses into two independent, continuous timeseries. Then, to evaluate whether the immediate and retrospective timeseries were representative of the same underlying fluctuations in attentional engagement, we assessed for the degree of overlap between the two (c) at a time lag of 0 ( $r_0$ ) and (d) at various other time lags (e.g., +1 minute;  $r_{+1}$ ).

immediate and retrospective timeseries when the two directly overlap, at a time lag of 0 (Figure 1c). Then, we assess cross-correlation scores between the two when the retrospective timeseries is shifted forwards or backwards in time, at various other time lags (e.g., at a time lag of +1 minute, as in Figure 1d). Given that cross-correlation scores can range from +1 (perfectly correlated) to 0 (not correlated) to -1 (negatively correlated), if immediate and retrospective timeseries represent the same underlying fluctuations in attentional engagement, we would expect to find a maximally-positive cross-correlation score when the timeseries directly overlap at a time lag of 0 compared to when the timeseries are shifted in some manner at any other time lag (as in Figure 2).

**Figure 2.**

*A Sample Participant's Cross-Correlations for the Less-Engaging Video Condition.*



*Note:* Within the sample participant's less-engaging data, cross-correlation scores were maximal at a time lag of 0 ( $r_0$ ), indicating a high degree of overlap in attentional-engagement fluctuations between the immediate and retrospective timeseries.

## Method

### *Participants*

It was determined *a priori* that 103 participants would be required for within-subject comparisons<sup>1</sup>. A strategic overshoot of this sample size was planned to account for participants being recruited in groups and for expected attrition due to quality control in online testing (Peer et al., 2014). In total, 118 participants (89 women, 26 men, 3 unknown;  $M_{age} = 21.2$  years,  $SD_{age} = 8.3$  years) were recruited for the experiment through the University of Waterloo's online participant pool. Informed written consent was obtained from all participants, and they received partial course credit as compensation for their time. All protocols and procedures were approved by the university's research ethics board.

Prior to analysis, we removed data from any participants who indicated that they had previously viewed at least one of the video lectures used for the study given that prior knowledge of the video content could impact both immediate and retrospective ratings of attentional engagement (13 participants). We also removed data from any participants that were non-responsive to over 20% of immediate or retrospective thought probes given the difficulty of tracking attentional engagement accurately in these cases (3 participants). As such, data from 102 participants (82 women, 19 men, 1 unknown;  $M_{age} = 21.5$  years,  $SD_{age} = 8.9$  years) were included in the final analyses.

### *Materials & Stimuli*

**Video lectures.** Two 15-minute video lectures were used as the video stimuli for the task: an open-sourced Yale lecture on Roman architecture (Kleiner, 2009) and a TEDx Talk on 'The

---

<sup>1</sup> An *a-priori* power analysis was conducted using G\*Power (test family = exact, statistical test = correlation; Faul et al., 2009) to determine the required sample size needed to detect a small effect size (Cohen et al., 1988). Based on an alpha of .05 and power of .95, it was determined that 103 participants would be needed.

Super Mario Effect' (Rober, 2018). The videos contained both audio and video elements, and each video was displayed at full screen resolution, embedded in the browser, with no playback controls visible. The TEDx Talk was hypothesized to be more engaging than the Yale lecture since the former involved varying camera angles, an animated presenter, colourful slides, and a more captivating topic. Prior to the task, each participant was asked to report on their familiarity with the video topics on a scale from 1 (not at all familiar) to 7 (very familiar). Overall, participants did not report any differences in familiarity between the two topics (art history,  $M = 2.3$ ,  $SD = 1.4$ ; computer science  $M = 2.7$ ,  $SD = 1.8$ ;  $t[101] = 1.76$ ,  $p = .081$ ,  $d_z = .17$ ).

**Immediate thought probes.** Participants' in-the-moment attentional engagement was assessed for each video lecture using 10 immediate thought probes (i.e., 20 probes in total for both videos) that appeared every 1-2 minutes throughout each lecture. Each immediate probe asked, "*Just before this screen appeared, how engaged were you with this video lecture?*", with response options presented as a multiple-choice Likert-style scale ranging from 1 (not at all) to 7 (very) as black text on a grey background.

**Retrospective thought probes.** Participants' retrospective attentional engagement was measured using 10 retrospective thought probes (i.e., 20 probes in total for both videos) that involved presenting participants with a 15-second video clip from the original video followed by a thought probe (i.e., video-stimulated recall). Similar to the immediate thought probes, each retrospective probe asked, "*How engaged were you when you first watched this section of the video lecture?*", with response options presented as a multiple-choice Likert-style scale ranging from 1 (not at all) to 7 (very) as black text on a grey background.

All retrospective thought-probe video clips were presented in sequential runtime order, that is, in the same temporal order that they originally appeared in the video. In addition,

retrospective probe timestamps were programmed to occur at least 10 seconds outside of any immediate probe timestamps, such that retrospective probes never overlapped with any immediate probes for the same participant, i.e., participants' attentional engagement was never probed twice for the same content. This was done to ensure that retrospective responses were not based on participants' initial immediate responses, and instead based on their actual RMAS.

**Post-task quiz.** Participants completed a multiple-choice quiz following each video lecture. The two quizzes each contained five items pertaining to the content of the video lecture they had just viewed with three or four response options each. The purpose of these quizzes was (1) to motivate participants to pay attention to the videos, and (2) to have an additional index of attentional engagement throughout the videos (e.g., if a participant both indicated that they were highly engaged at the end of the video and answered the content question from that section correctly). A copy of the quizzes used can be found in Appendix A.

**Other materials.** There were two additional measures employed within the design that were not directly related to the aims of the current study and are thus not included in the current analyses. First, following the completion of the video-viewing tasks including all immediate and retrospective thought probes, there was a self-reported measure of participants' level of motivation for completing the study's task well. Then, there was a self-reported measure of participants' general attentional-control tendencies (Derryberry & Reed, 2002).

### ***Procedure***

To begin, participants were first presented with a letter of informed consent before being asked to provide their demographic information (gender, age), to rate their familiarity with the two videos' overarching topics, and to rate their level of motivation for performing well on the task. Participants then received the instructions to complete the video-viewing task, which

stipulated that they would watch a video lecture, responding to intermediate thought probes throughout, and that there would be a test on their knowledge of the video content afterwards. Additionally, participants were asked not to take notes or multitask while viewing the lecture.

The main task components of the study are shown in Figure 3. Participants watched the video lectures, and their in-the-moment levels of attentional engagement were measured using immediate thought probes. Then, after viewing both video lectures, they completed the video-stimulated recall component of the study. During this component, participants were presented with 15-second clips from each video they had just viewed. For each clip they were asked to retrospectively report their level of engagement the first time they watched that section of the video. Then, participants completed the recall quizzes for both videos. The order of the two video conditions was counterbalanced between participants. Finally, participants completed the self-report measure of attentional control (not analyzed here).

## **Results**

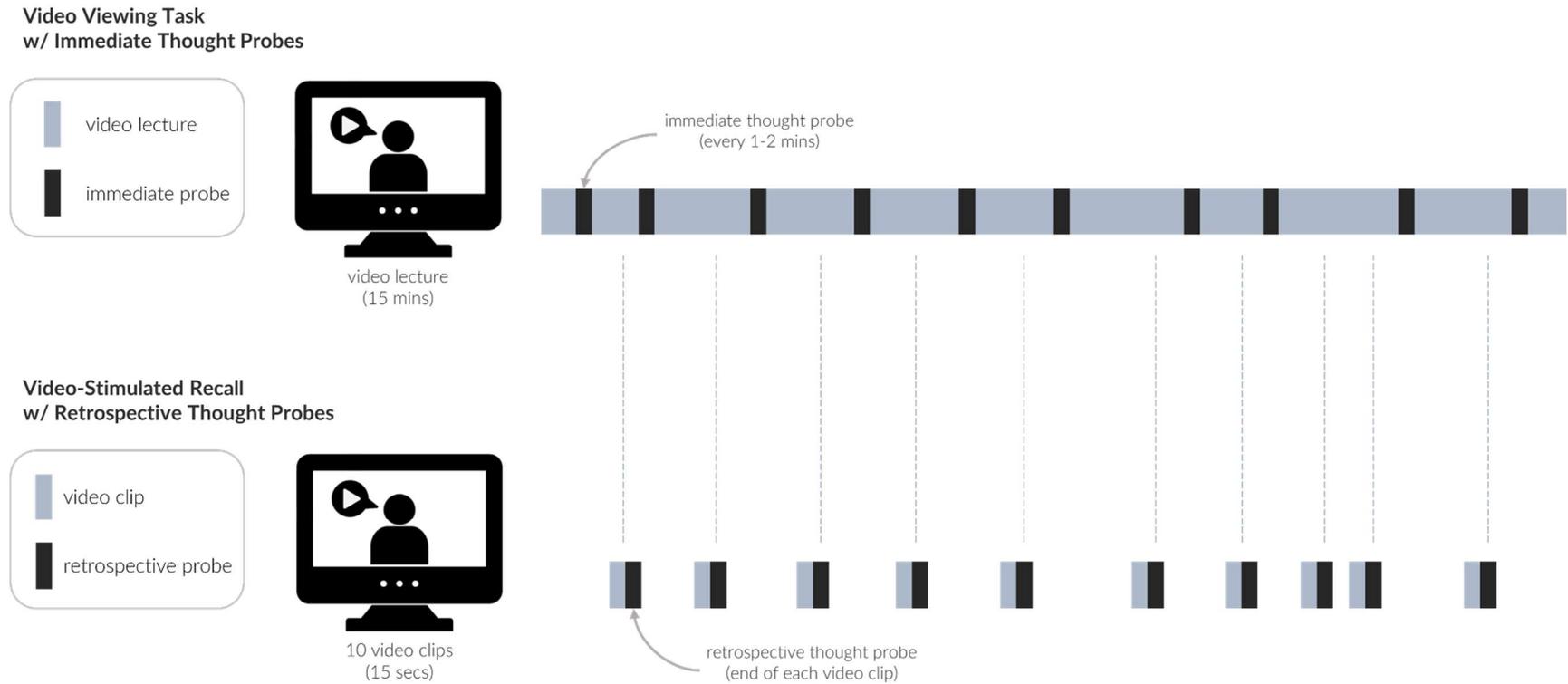
The key dependent variable assessed in this experiment was participants' self-reported attentional engagement. Additionally, recall of video content was assessed as a marker of memory for the two video conditions. The two key independent variables in this experiment were probe type (immediate versus retrospective) and video type (less-engaging versus more-engaging). All measures were normally distributed (skew <3 and kurtosis <10; Kline, 1998).

### ***Probe Type and Video Type***

A two-way repeated-measures ANOVA was conducted to assess the effects of probe type (immediate versus retrospective) and video type (less-engaging versus more-engaging) on participants' self-reported attentional engagement scores (see Figure 4). There was not a

**Figure 3.**

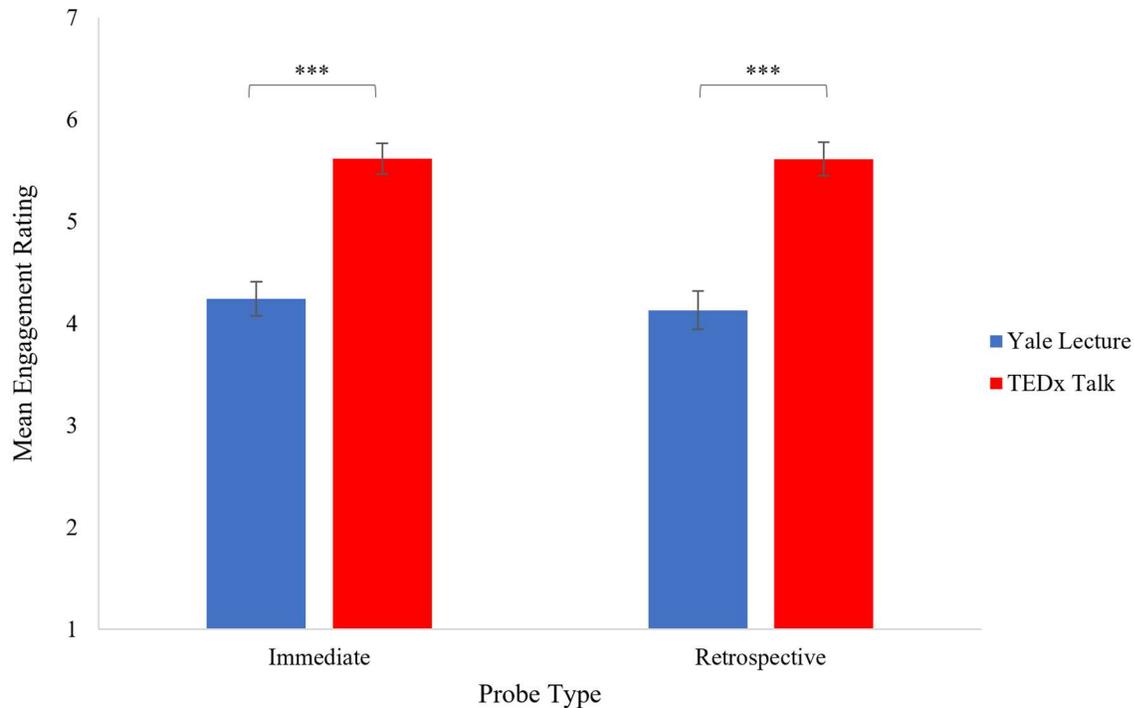
*The Video-Viewing Task and Video-Stimulated Recall Task Procedures for Experiment 1.*



*Note.* Dotted lines indicate that the immediate and retrospective probes never occurred at the same time points in the video for the same participant. I.e., participants were never probed on the same section of the video lectures twice. Participants completed the video-viewing task for both video conditions before then completing the video-stimulated recall task for both video conditions.

**Figure 4.**

*Mean Engagement Probe Responses for the ‘Less-Engaging’ Yale Lecture and ‘More-Engaging’ TEDx Talk Video Conditions as a Function of Immediate and Retrospective Probe Types (N = 102).*



*Note:* Error bars represent +/- 1 standard error of the mean (SEM).

\*\*\* Significant difference at the .001 level.

significant interaction between the two independent variables ( $F[1, 101] = 1.35, p = .248, \eta^2_p = .01$ ), nor was there a significant main effect of probe type on engagement scores ( $F[1, 101] = .51, p = .475, \eta^2_p = .01$ ). This indicated that participants had similar overall ratings of attentional engagement immediately during the video-viewing task and retrospectively. However, there was a significant main effect of video type ( $F[1, 101] = 74.36, p < .001, \eta^2_p = .42$ ), with participants' overall engagement ratings being significantly lower for the Yale lecture than for the TEDx Talk. This supported defining the Yale lecture as the 'less-engaging' video condition and the TEDx Talk as the 'more-engaging' video condition.

### ***Recall of Video Content***

On average, participants scored significantly higher on the post-task quiz for the TEDx Talk ( $M = 3.4$  [68%],  $SD = 0.9$ ) than for the Yale lecture ( $M = 3.1$  [62%],  $SD = 1.2$ ;  $t[101] = 2.92$ ,  $p = .004$ ,  $d_z = .29$ ). As seen in Table 1, participants' immediate and retrospective probe responses were strongly correlated for both video conditions. Furthermore, participants' immediate and retrospective probe responses for attentional engagement were positively correlated with their post-task quiz responses, suggesting that participants who reported higher attentional engagement both in the moment and retrospectively also tended to perform better on the recall quiz than participants who reported lower attentional engagement. Although the patterns were similar in the less-engaging and more-engaging video conditions, the immediate reports of attentional engagement were more strongly correlated with the recall quiz results than were the retrospective reports of attentional engagement. For these comparisons, we applied a calculation for the test of the difference between two dependent correlations with one variable in common (Lee & Preacher, 2013, based on Steiger, 1980). For the less-engaging video condition, the immediate probes correlated more strongly with the quiz results ( $r = .43$ ) than did the retrospective probes ( $r = .32$ , one-tailed  $p = .005$ ). Similarly, for the more-engaging video condition, the immediate probes correlated more strongly with the quiz results ( $r = .43$ ) than did the retrospective probes ( $r = .28$ , one-tailed  $p = .004$ ).

### ***Cross-Correlation Results***

Using the cross-correlation analytic approach, we applied the following steps to determine the temporal overlap between immediate and retrospective thought probes. First, for each video condition, we converted our immediate and retrospective probe responses into two separate, continuous timeseries by interpolating each set of probe responses at 1-minute intervals

**Table 1.**

*Correlations Between Attentional-Engagement Probe Responses and Content-Recall Quiz Results for Experiment 1 (N = 102).*

	1	2	3
<b>Less-Engaging Video</b>			
1. Immediate Probe		.89***	.43***
2. Retrospective Probe			.32**
3. Recall Quiz			
<b>More-Engaging Video</b>			
1. Immediate Probe		.81***	.43***
2. Retrospective Probe			.28**
3. Recall Quiz			

\*\*\* Significant at the .001 level.

\*\* Significant at the .01 level.

using a piecewise polynomial function (i.e., PCHIP, Piecewise Cubic Hermite Interpolating Polynomial; Fritsch & Carlson, 1980; Kahaner et al., 1988). This function was selected due to its ability to interpolate data within the bounds of the data points to best approximate dynamic and fluctuating timeseries (Barker & McDougall, 2020; Dan et al., 2020; Zeinali et al., 2014). This resulted in two 16-point timeseries that best preserved the original probe responses. Then, for each video condition, we calculated cross-correlation scores between the immediate and retrospective timeseries using time steps of 1 minute. This step size was chosen to reflect the interpolation intervals used, resulting in 19 cross-correlation time lag scores ranging from -9-minute to +9-minutes for each participant per video condition.

Next, we used the following stepwise analysis plan to comprehensively assess the cross-correlation results: to evaluate the concordance between the immediate and retrospective reports, we (1) determined whether the overlap between the two timeseries was significantly different from a null effect, using one-sample t-tests (compared to 0) to assess mean cross-correlation scores at a time lag of 0; then (2) determined if the cross-correlation coefficients at a time lag of

0 were the peakiest<sup>2</sup> by comparing the 95% confidence intervals surrounding each of the time lags visually; finally, to determine if any differences existed between video conditions, we (3) used a paired-samples t-test to compare mean cross-correlation scores at a time lag of 0 between the less-engaging and more-engaging video conditions.

As in typical Pearson correlations, a cross-correlation coefficient cannot be computed if there is zero variability within a participant's data. Therefore, participants who had zero variability within either their immediate or retrospective probe responses (e.g., responded to all thought probes with '1') were deleted listwise, resulting in 63 participants' cross-correlation data being included in our comparisons.<sup>3</sup>

**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients of attentional engagement significantly differed from 0 for both the less-engaging ( $M = .32, SD = .44$ ; one-sample  $t[62] = 5.85, p < .001, d_z = .74$ ; Figure 5a) and more-engaging video conditions ( $M = .26, SD = .39$ ; one-sample  $t[62] = 5.21, p < .001, d_z = .66$ ; Figure 5b). Therefore, participants rated their attentional engagement similarly when probed immediately and retrospectively, demonstrating that specific prior moments of attentional engagement can be accurately probed after the fact.

The 95% confidence intervals in Figures 5a and 5b indicate that the cross-correlation coefficients were highest around the 0 lag for both video conditions, as they were significantly higher than time lags -9 to -2 and +3 to +9. Therefore, the congruence between participants' immediate and retrospective reports of attentional engagement was strongest when the timeseries

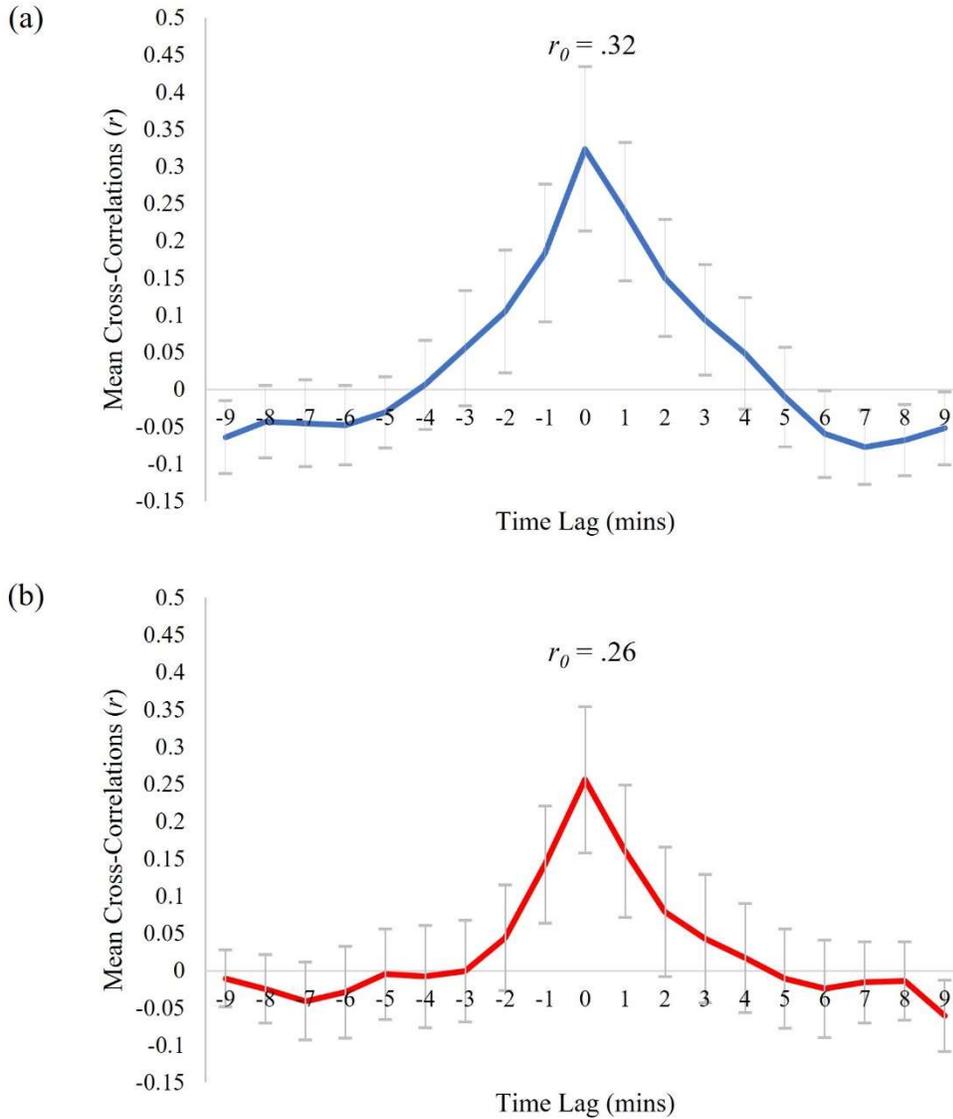
---

<sup>2</sup> A 'peakier' cross-correlation = high cross-correlation coefficient at a time lag of 0 with other time lags dropping off on either side; a less 'peaky' cross-correlation = low cross-correlation coefficient at a time lag of 0, with or without other time lags dropping off on either side.

<sup>3</sup> This reduction of sample size did not impact the results of the analyses compared to when all possible participants were included (see Appendix B).

**Figure 5.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging and (b) More-Engaging Video Conditions of Experiment 1 (N = 63).



Note. Error bars represent 95% confidence intervals.

were directly overlaid and/or only shifted by up to 2 minutes, suggesting that participants were reporting on their RMASs of attentional engagement.

**Comparing video conditions.** Finally, the cross-correlation coefficients for the less-engaging and more-engaging video conditions did not differ significantly at a time lag of 0 ( $t[62]$

= .97,  $p = .337$ ,  $d_z = .16$ ). Therefore, the video condition did not affect participants' concordance between the immediate and retrospective reports.

## **Discussion**

The goal of Experiment 1 was to examine the overarching RMAS of *attentional engagement* by assessing the degree of correspondence between individuals' immediate and retrospective probe reports of this attentional state. Our findings indicated that when using short cues from the original video-viewing task, participants' reports of memories of attentional states corresponded moderately well with their in-the-moment reports. This was observed both when looking at the aggregated reports and within the temporal dynamics of the probes across the task. Furthermore, our findings indicated that this correspondence was not affected by the overall engagingness of the videos used for the task. Therefore, at least when probing about individuals' attentional engagement, retrospective reports of RMASs can be similar to in-the-moment reports of an initial attentional experience.

## Experiment 2: Mind-Wandering

In Experiment 2, we examined the more specific RMAS of *mind-wandering* by assessing the degree of correspondence between individuals' immediate and retrospective probe reports of this attentional state. When one is in a state of mind-wandering, attention drifts from the primary task to unrelated thoughts (Smallwood & Schooler, 2006). Although it might be assumed that attentional engagement and mind-wandering occur on opposite ends of an attentional spectrum, this has been called into question (Anderson, 2011; Seli et al., 2018; Weinstein et al., 2018). Instead, it has been suggested that mind-wandering is only one form attentional *disengagement* (Danckert, 2018), with other forms of disengagement including states such as external distractions (Stawarczyk et al., 2011) and blanking (Ward & Wegner, 2013). Furthermore, because it has been shown that the type of state being probed might affect how participants' reports on their attentional experiences (e.g., "*At the time of the ding, my mind was on something other than the text*" versus "*At the time of the ding, my mind was on the text*"; Weinstein et al., 2018), one's degree of mind-wandering cannot be assumed based on their report of attentional engagement or vice-versa (Weinstein et al., 2018). Thus, it is unclear whether RMASs for mind-wandering will show similar correspondence with immediate reports to what we found when assessing RMASs for overall attentional engagement.

Mind-wandering experiences are often divided into two separate classifications: intentional mind-wandering, wherein an individual has deliberately let their thoughts drift away from the task at hand; and unintentional mind-wandering, wherein an individual might find that their thoughts have drifted away from the task without their explicit control (Giambra, 1995; Seli, Risko, & Smilek, 2016; Seli, Risko, Smilek, et al., 2016). Although individuals might experience both types of mind-wandering throughout a task, it is important to differentiate

between the two. This is because intentional and unintentional mind-wandering have been found to: (a) occur at different rates depending on context, such as intentional mind-wandering occurring more often in an easier attention task and unintentional mind-wandering occurring more often in a more difficult attention task (Seli, Risko, & Smilek, 2016); and (b) to associate with different behaviours, such as intentional but not unintentional mind-wandering relating to fidgeting (Carriere et al., 2013). Therefore, in the present study we distinguished between intentional and unintentional mind-wandering to understand how either might be affected in retrospective reporting.

In Experiment 2, we examined individuals' retrospective reporting of intentional and unintentional mind-wandering for the same two video lectures as in Experiment 1. We had participants complete the same short quizzes at the end of each video condition to assess their recall for the video content, and we assessed the concordance between the temporal dynamics of participants' immediate and retrospective reports of mind-wandering for both video conditions. Our analyses were aimed at addressing similar issues as in Experiment 1.

## **Method**

### ***Participants***

One hundred and twenty-three participants (87 women, 31 men, 1 non-binary, 3 unknown;  $M_{age} = 19.7$  years,  $SD_{age} = 3.1$  years) who had not participated in Experiment 1 were recruited for the experiment through the University of Waterloo's online participant pool. Informed written consent was obtained from all participants, and they received partial course credit as compensation for their time. All protocols and procedures were approved by the university's research ethics board.

Prior to data analysis, we removed data from any participants who indicated that they had previously viewed at least one of the video lectures used for the study (19 participants) or if they were non-responsive to over 20% of immediate or retrospective thought probes (7 participants). As such, data from 97 participants (73 women, 20 men, 1 non-binary, 3 unknown;  $M_{age} = 19.9$  years,  $SD_{age} = 3.4$  years) were included in the final analyses.

### ***Materials & Stimuli***

The materials and stimuli were identical to those used in Experiment 1, except the immediate and retrospective thought probes asked about participants' subjective reports of intentional and unintentional mind wandering on a scale from 0 (not at all) to 10 (completely). Immediate probe: *“To what extent were you [intentionally/unintentionally] mind wandering just before this probe appeared?”*. Retrospective probe: *“To what extent were you [intentionally/unintentionally] mind wandering when you first watched this section of the video lecture?”*. Twenty immediate probes were presented to each participant (10 intentional, 10 unintentional) and 10 retrospective clips were shown to each participant (each with an intentional and unintentional probe, counterbalanced). Overall, participants did not report any differences in familiarity between the two topics (computer science,  $M = 2.1$ ,  $SD = 1.6$ ; art history,  $M = 2.0$ ,  $SD = 1.2$ ;  $t[96] = 0.62$ ,  $p = .537$ ,  $d_z = .06$ ).

### ***Procedure***

The procedures were identical to those of Experiment 1, except for the probe wording (mind-wandering instead of attentional engagement) and the addition of the independent variable of mind-wandering intentionality (intentional versus unintentional).

## Results

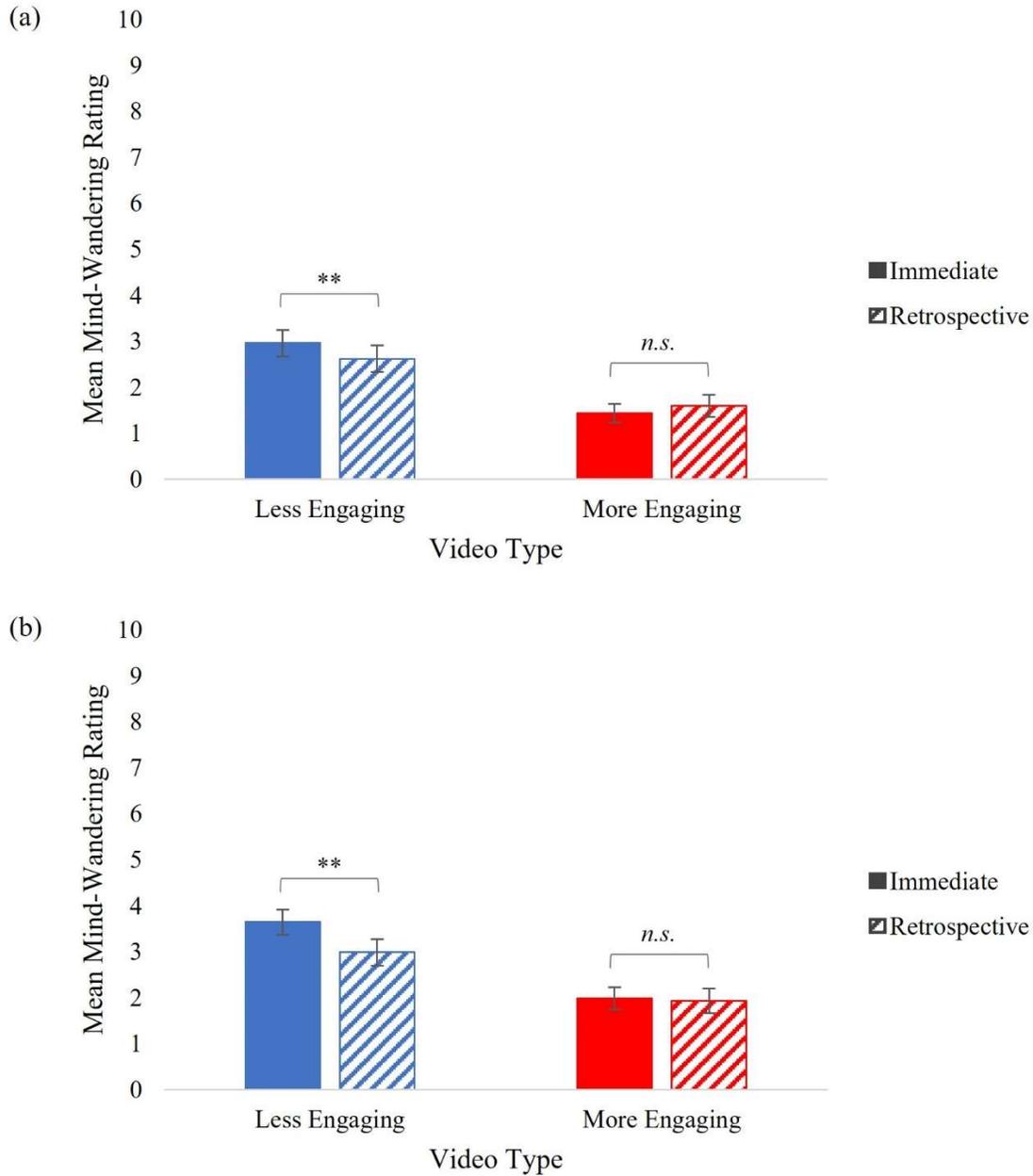
The key dependent variable assessed in this experiment was participants' self-reported mind-wandering. Additionally, recall of video content was assessed as a marker of memory. The three key independent variables in this experiment were probe type (immediate versus retrospective), video type (less-engaging versus more-engaging), and mind-wandering intentionality (intentional versus unintentional). All measures were normally distributed (skew  $<3$  and kurtosis  $<10$ ; Kline, 1998).

### *Probe Type, Video Type, and Intentionality*

A three-way repeated-measures ANOVA was conducted to assess the effects of probe type (immediate versus retrospective), video type (less-engaging versus more-engaging), and mind-wandering intentionality (intentional versus unintentional) on participants' self-reported mind-wandering scores (see Figure 6). There was not a significant main effect of probe type ( $F[1, 96] = 2.11, p = .150, \eta^2_p = .02$ ). However, there was a significant main effect of video type ( $F[1, 96] = 78.08, p < .001, \eta^2_p = .45$ ) with participants reporting significantly higher mind-wandering in the less-engaging video condition than in the more-engaging video condition. This effect was qualified by a significant two-way interaction with probe type ( $F[1, 96] = 7.60, p = .007, \eta^2_p = .07$ ), indicating that the immediate probe responses were significantly higher than the retrospective probe responses in the less-engaging video condition but that there was no significant difference between probe types in the more-engaging video condition. Additionally, there was a significant main effect of mind-wandering intentionality ( $F[1, 96] = 9.89, p = .002, \eta^2_p = .09$ ) with participants reporting significantly higher unintentional mind-wandering than intentional mind-wandering overall. This effect was not qualified by either a two-way interaction with probe type ( $F[1, 96] = 2.57, p = .112, \eta^2_p = .03$ ) nor with video type ( $F[1, 96] = .25, p =$

**Figure 6.**

Mean Immediate and Retrospective (a) Intentional and (b) Unintentional Mind-Wandering Probe Responses for the Less-Engaging and More-Engaging Video Conditions of Experiment 2 (N = 97).



Note: Error bars represent +/- 1 standard error of the mean (SEM).

\*\* Significant difference at the .01 level.

n.s. Not a significant difference at the .05 level.

.619,  $\eta^2_p = .00$ ). Finally, the three-way interaction between the three independent variables did not reach significance ( $F[1, 96] = .15, p = .704, \eta^2_p = .00$ ).

Therefore, participants reported significantly less mind-wandering retrospectively than they had in the moment, suggesting that perhaps specific mind-wandering experiences aren't as memorable or distinguishable in less-engaging contexts such as the Yale lecture video. Furthermore, participants reported mind-wandering unintentionally more than intentionally overall, which has been noted in some prior studies (e.g., Seli, Risko, & Smilek, 2016) but not in others (e.g., Wammes, Boucher, et al., 2016).

### ***Recall of Video Content***

As with Experiment 1, participants scored significantly higher on the TEDx Talk's post-task quiz ( $M = 3.5$  [70%],  $SD = 0.8$ ) than on the Yale lecture's post-task quiz ( $M = 3.1$  [62%],  $SD = 1.3$ ;  $t[96] = 2.96, p = .004, d_z = .30$ ). As seen in Table 2, participants' immediate and retrospective probe responses for mind-wandering were negatively correlated with their post-task quiz responses, suggesting that participants who reported more mind-wandering both in the moment and retrospectively also tended to perform poorer on the recall quiz than participants who reported less mind-wandering. Unlike in Experiment 1, the correlations between the content recall and the immediate reports of mind-wandering were not statistically different from the correlations between content recall and the retrospective reports of mind-wandering for either video condition. For these comparisons, we again tested the difference between two dependent correlations with one variable in common (Lee & Preacher, 2013, based on Steiger, 1980). For the less-engaging video condition, the correlations between the immediate probes and the quiz results were similar to the correlations between the retrospective probes and the quiz results for both the intentional probes (intentional immediate  $r = -.48$  and retrospective  $r = -.47$ , one-tailed

**Table 2.**

*Correlations Between Mind-Wandering Probe Responses and Content-Recall Quiz Results for Experiment 2 (N = 97).*

	1	2	3	4	5
<b>Less-Engaging Video</b>					
<i>Intentional MW</i>					
1. Immediate Probe		.69***	.61***	.55***	-.48***
2. Retrospective Probe			.51***	.68***	-.47***
<i>Unintentional MW</i>					
3. Immediate Probe				.69***	-.39***
4. Retrospective Probe					-.36***
<i>Content Recall</i>					
5. Recall Quiz					
<b>More-Engaging Video</b>					
<i>Intentional MW</i>					
1. Immediate Probe		.66***	.81***	.53***	-.40***
2. Retrospective Probe			.56***	.77***	-.32**
<i>Unintentional MW</i>					
3. Immediate Probe				.69***	-.35***
4. Retrospective Probe					-.23*
<i>Content Recall</i>					
5. Recall Quiz					

*Note.* MW = mind-wandering.

\*\*\* Significant at the .001 level.

\*\* Significant at the .01 level.

\* Significant at the .05 level.

$p = .4$ ; unintentional immediate  $r = -.39$  and retrospective  $r = -.36$ , one-tailed  $p = .3$ ). The same was observed for the more-engaging video condition (intentional immediate  $r = -.40$  and retrospective  $r = -.32$ , one-tailed  $p = .2$ ; unintentional immediate  $r = -.35$  and retrospective  $r = -.23$ , one-tailed  $p = .1$ ).

### **Cross-Correlation Results**

Cross-correlation calculations were identical to those of Experiment 1, with the addition of the independent variable of mind-wandering intentionality (intentional versus unintentional).

We then used the same stepwise analysis plan to comprehensively assess whether the immediate and retrospective timeseries for each type of mind-wandering probe in each video lecture were representative of the same underlying fluctuations in mind-wandering.

Participants who had zero variability within either their immediate or retrospective probe reports (e.g., responded to all thought probes with '1') were deleted listwise, resulting in 58 participants' cross-correlation data being included in our comparisons.<sup>4</sup>

**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients only differed significantly from 0 for the less-engaging video's intentional probes ( $M = .12$ ,  $SD = .38$ ; one-sample  $t[57] = 2.47$ ,  $p = .016$ ,  $d_z = .32$ ; Figure 7a) but not for the less-engaging video's unintentional probes ( $M = .02$ ,  $SD = .40$ ; one-sample  $t[57] = .47$ ,  $p = .644$ ,  $d_z = .06$ ; Figure 7b). Furthermore, the average cross-correlation coefficients for the more-engaging video condition did not differ significantly from 0 for either the intentional probes ( $M = .01$ ,  $SD = .32$ ; one-sample  $t[57] = .29$ ,  $p = .774$ ,  $d_z = .04$ ; Figure 7c) nor the unintentional probes ( $M = -.04$ ,  $SD = .38$ ; one-sample  $t[57] = -.86$ ,  $p = .395$ ,  $d_z = -.11$ ; Figure 7d). Therefore, participants only rated their mind-wandering similarly when probed immediately and retrospectively about their intentional mind-wandering during one video condition, demonstrating that mind-wandering might not be accurately probed after the fact.

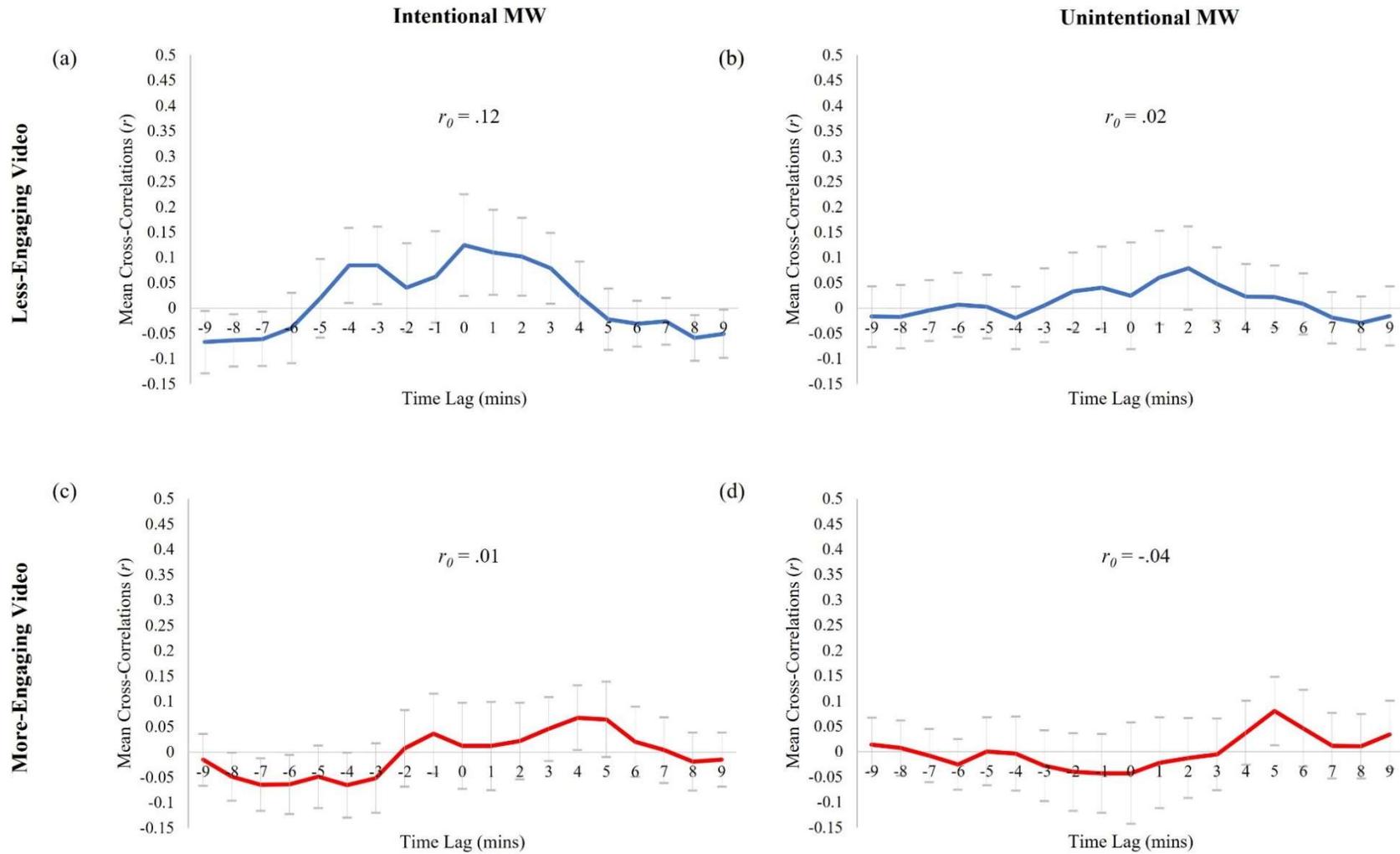
The 95% confidence intervals in Figures 7a-d suggest that the cross-correlation coefficients at the 0 lag were not significantly higher than at any other time lag, except for the less-engaging intentional mind-wandering condition (Figure 7a) where the 0 lag was significantly higher than lags -7 to -9 and +6 to +9. Therefore, the overall congruence between participants' immediate and retrospective reports of mind-wandering were not strongest when

---

<sup>4</sup> This reduction of sample size did not impact the results of the analyses compared to when all possible participants were included (see Appendix C).

**Figure 7.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging Intentional MW, (b) Less-Engaging Unintentional MW, (c) More-Engaging Intentional MW, & (d) More-Engaging Unintentional MW Conditions of Experiment 2 (N = 58).



Note. MW = mind-wandering. Error bars represent 95% confidence intervals.

directly overlaid, suggesting that participants might not have been reporting on their RMASs of mind-wandering or that they had poor recollection of those states.

**Comparing video conditions.** Finally, there was not a significant interaction between video type (less-engaging versus more-engaging) and mind-wandering type (intentional versus unintentional) on mean cross-correlation coefficients at a time lag of 0 ( $F[1, 57] = .29, p = .594, \eta^2_p = .01$ ). There was also not a significant main effect of either video type ( $F[1, 57] = 2.86, p = .096, \eta^2_p = .05$ ) nor of mind-wandering type ( $F[1, 57] = 2.90, p = .094, \eta^2_p = .05$ ). Therefore, neither video condition nor mind-wandering type affected participants' concordance between the immediate and retrospective reports.

## Discussion

The goal of Experiment 2 was to examine the more specific RMAS of *mind-wandering* by assessing the degree of correspondence between individuals' immediate and retrospective probe reports of this attentional state. Our findings indicate that probing mind-wandering retrospectively with short cues from the original video-viewing task does not always result in a correspondence with immediate probes collected during the initial task. When probes were aggregated into an average score, we observed that participants reported less mind-wandering retrospectively than they had in the moment. This was only found for the less-engaging video condition, suggesting that there was a discrepancy between how they experienced mind-wandering during the original task and how they reported on that mind-wandering after the fact. This finding was also observed within the temporal dynamics of the probes across the task, with there being low correspondence between immediate and retrospective probe reports of intentional mind-wandering in the less-engaging video condition and no correspondence for the rest of the comparisons (unintentional mind-wandering in either video condition and intentional

mind-wandering in the more-engaging video condition), although the differences between conditions and intentionality were not significant overall. Therefore, when probing about individuals' mind-wandering, retrospective reports of RMASs are not always similar to in-the-moment reports of an initial attentional experience.

## Comparing Attentional Engagement and Mind-Wandering

Although it might appear that the results from Experiment 1 and Experiment 2 were drastically different with momentary states of attentional engagement being recalled but mind-wandering being more nuanced, we aimed to directly compare the results of the two experiments to determine if these differences were significant. Specifically, we examined whether there was greater correspondence between immediate and retrospective probe reports when assessing attentional engagement than when assessing mind-wandering. To facilitate direct comparisons, we used combined intentional and unintentional mind-wandering results from Experiment 2<sup>5</sup> to compare attentional engagement with overall mind-wandering.

For this comparison, we standardized each participant's raw probe responses, re-ran the cross-correlation analyses, then compared the results using independent *t*-tests. For the less-engaging video condition, average cross-correlation coefficients were significantly higher at a time lag of 0 when asking participants about attentional engagement ( $M = .31, SD = .44$ ; Figure 8a) than when asking about overall mind-wandering ( $M = .11, SD = .42$ ; Figure 8b;  $t[119] = 2.62, p = .010, d = .48$ ). The same was observed for the more-engaging video condition: average cross-correlation coefficients were significantly higher at a time lag of 0 for attentional engagement ( $M = .25, SD = .41$ ; Figure 8c) than for overall mind-wandering ( $M = .08, SD = .42$ ; Figure 8d;  $t[119] = 2.14, p = .034, d = .39$ ). Therefore, there was stronger concordance between the immediate and retrospective reports of attention when participants were probed about levels of attentional engagement compared to when they were probed about levels of mind-wandering.<sup>6</sup>

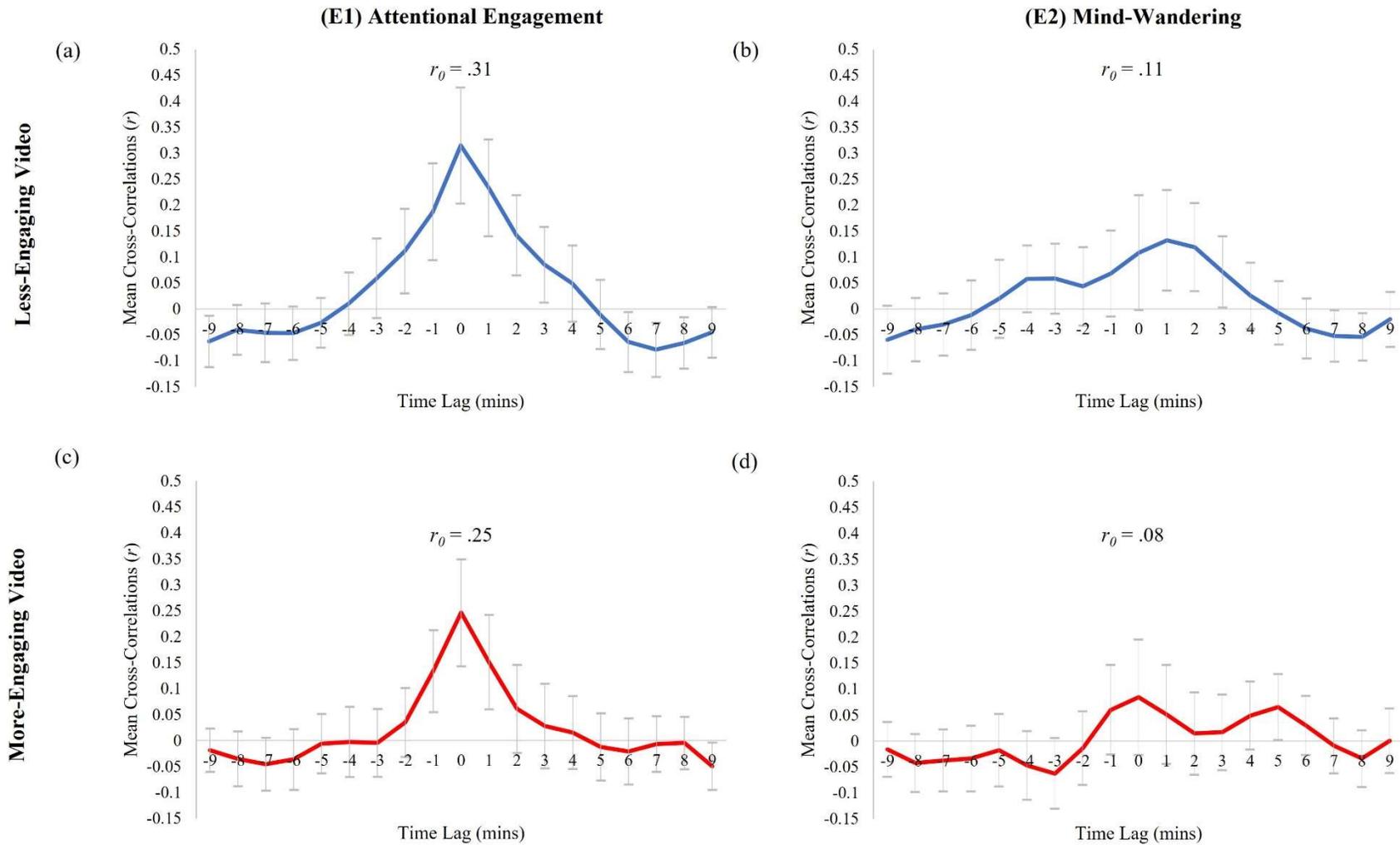
---

<sup>5</sup> To investigate overall mind-wandering reports, we averaged each participant's raw intentional and unintentional mind-wandering scores into overall mind-wandering scores. As a sanity check, we also re-ran the cross-correlation analyses on these combined scores, which had similar results as when intentional and unintentional mind-wandering cross-correlations were examined separately (see Appendix D).

<sup>6</sup> The same results were observed when including all possible participants in the analyses (see Appendix E).

**Figure 8.**

Mean Cross-Correlation Scores as a Function of Time Lag from the Standardized (a) E1 Less-Engaging, (b) E2 Less-Engaging, (c) E1 More-Engaging and (d) E2 More-Engaging Conditions ( $N_{E1} = 63, N_{E2} = 58$ ).



Note. E1 = Experiment 1 (attentional-engagement probes), E2 = Experiment 2 (overall mind-wandering probes). Error bars represent 95% confidence intervals.

## General Discussion

Across two experiments, we examined the degree of correspondence between the temporal dynamics of immediately and retrospectively probed self-reports of attention. In Experiment 1, when participants reported on their remembered momentary attentional states (RMASs) of engagement, we found moderate correspondence between in-the-moment (immediate) reports and their RMASs in two different video-lecture conditions. This correspondence was found both when the probes were aggregated into an average score and when the timeseries of the probes were compared, supporting the conclusion that individuals could recall specific momentary experiences of attentional engagement. However, the results were not so clear in Experiment 2, in which participants reported on the more specific RMAS of mind-wandering (intention and unintentional). When examining the aggregated levels of this attentional state, we found that participants rated their mind-wandering (intentional and unintentional) similarly across in-the-moment and retrospective probes for the more-engaging TEDx Talk but that they reported less mind-wandering retrospectively than they had in-the-moment for the less-engaging Yale lecture. When the timeseries of the probes were compared, only intentional mind-wandering reports for the less-engaging video showed a significant degree of correspondence between retrospective and immediate probes.

When the results of the two experiments were compared directly, it became clear that there was significantly less temporal correspondence between participants' immediate and retrospective reports of mind-wandering (Experiment 2) than of attentional engagement (Experiment 1). The results of this exploratory between-study comparison suggest that perhaps a broader RMAS such as attentional engagement displays stronger correspondence between the temporal dynamics of immediate and retrospective probe reports than a more specific RMAS

such as mind-wandering. Taken together, these results are consistent with the notion that people have enduring memories for momentary episodes of general attentional engagement, but a much weaker memory for their momentary episodes of the more specific state of mind-wandering.

While we suggest that the degree of correspondence between immediate and retrospective probes for attentional engagement is based on enduring memory for prior moments of engagement, it is possible that other factors might have contributed to participants' reports of their RMASSs. One alternative explanation to the present findings could be that individuals do not remember their initial momentary attentional state and instead base retrospective reports on their memory for the material covered in the video cues presented at recall. That is, when participants are shown the video clip to cue their recall, they may first evaluate if they remember the content of the clip and then report an RMASS consistent with the success of the recall of the video clip. Indeed, it has been suggested that perhaps individuals make inferences regarding their recent attentive states based on their recall for immediately preceding task events. Specifically, Dixon and Li (2013) hypothesized that individuals infer whether they have been mind-wandering just prior to an immediate attention probe by assessing if their working memory contains 'detailed and elaborate information' about the prior task content. A related concern is that a state of inattentiveness might preclude someone from accurately assessing their attentional state. However, there is evidence that individuals can be aware of their own inattentiveness as people are able to self-catch their episodes of mind-wandering (e.g., Varao-Sousa & Kingstone, 2019). Given these concerns, future research might focus on investigating how individuals make inferences about their prior attentive states; perhaps people use various cues to infer a RMASS, such as cues from the retrospective stimulus itself.

A noteworthy aspect of our study is that we did not find a strong concordance between immediate and retrospective probes for mind-wandering, an often-studied attentional state (e.g., Banks et al., 2015; Barron et al., 2011; Kane et al., 2021; Smallwood & Schooler, 2006). One explanation for this outcome is that perhaps mind-wandering is not a very memorable experience. This would be in line with Kane et al.'s (2021) statement that they had “little confidence in retrospective reports of mind wandering” (p. 2402). However, an alternative explanation might be that people *do* have enduring memory for specific mind-wandering episodes but that we were not able to capture these memories successfully with our cued retrospective-report technique. Mind-wandering has been described as a ‘fleeting’ (Kane et al., 2021; Stawarczyk, 2018) and oftentimes ‘spontaneous’ experience (Carriere et al., 2013; Smallwood & Schooler, 2015), which suggests that levels of mind-wandering might change frequently – perhaps more frequently than overall engagement. As such, it is plausible that while participants viewed the video their mind-wandering states might have changed more rapidly than our probing method could sample. Because our immediate and retrospective probes sampled different timepoints of the task, the immediate and retrospective probes – even those proximate in the time series – might not have sampled the same momentary states. Future studies could investigate this possibility by employing stimulated-recall cues from the same or very similar timepoints as those of the immediate mind-wandering probes.

It is also worth noting that the cues implemented in our stimulated-recall task were 15-second clips from the previously viewed video lectures, which were then compared to immediate probes sampling the point right before participants were probed. We chose to use a 15-second clip to provide participants with sufficient contextual information to trigger their memory for their attentional state during that clip. However, this meant that the retrospective probes were

less temporally specific than the immediate probes. It is possible that the duration of the clips may have been too long to cue specific fleeting mind-wandering experiences, which may have impacted retrospective reports of mind-wandering. Future studies might explore implementing cues of varying durations (e.g., 5 seconds or still images) to assess whether individuals can effectively report on RMASSs with less contextual information; perhaps shorter cues might lead to more temporal specificity and still provide sufficient context to cue RMASSs.

### **Future Directions**

Given that people may be able to recall some types of RMASSs (attentional engagement) better than others (mind-wandering), future research might investigate what factors contribute to the degradation of these memories. Specifically, understanding whether prevalent theories of forgetting such as decay or interference (Demetriou & Bakracevic, 2009; Ricker et al., 2016) might apply to the stability of RMASSs could help clarify why some types of RMASSs are better recalled than others and at what delay individuals might begin forgetting their RMASSs. For example, a future study might investigate whether immediate and retrospective timeseries diverge with increased time between the initial task and the retrospective probes, which might indicate that memories of RMASSs degrade over time. Alternatively, future research might investigate whether immediate and retrospective timeseries are less similar in a higher memory-interference condition than in a lower memory-interference condition, which might indicate that having similar material within a task would increase interference with recall of RMASSs.

Future research could also examine people's retrospective reports of other momentary attentional states. Although the present studies only examined two attentional experiences (engagement and mind-wandering), future studies could investigate whether people have robust RMASSs for other states such as boredom (Danckert et al., 2018) or flow (Marty-Dugas et al.,

2021). Boredom is described in the literature as “a disengaged state in which the individual is motivated to be engaged with their environment, but for whom all attempts to do so fail” (Danckert et al., 2018, p. 24) and is primarily characterized as being an unpleasant experience. Considered from this perspective, it might be assumed that boredom is a salient negatively valenced attentional experience, which might lead to memorable RMASs. Perhaps individuals might recall short moments of reduced boredom during attempts to engage with the environment (e.g., switching on a television show or picking up a book), but then recall salient increases in boredom with each unsuccessful attempt at engagement. Alternatively, flow is often described as a 'peak' experience (Csikszentmihalyi & Csikszentmihalyi, 1992), therefore perhaps its saliency might allow for a particularly memorable RMAS. However, descriptions of flow experiences also typically involve a sense of timelessness and a lack of self-conscious awareness (Csikszentmihalyi & Csikszentmihalyi, 1992), suggesting that perhaps specific indices of flow might not be discernable in retrospect. Furthermore, it has been suggested that flow might unfold over long periods of time on task rather than in short bursts (Marty-Dugas et al., 2021), further suggesting that memories of specific momentary flow experiences might not be easily recalled with temporal specificity. Thus, exploring other types of RMASs is imperative to understanding how memory for attentional experiences might vary. Based on the current findings, it might be hypothesized that more general experiences can be better recalled than more specific experiences.

Finally, the retrospective-probing method employed in this paper could be used to capture temporally specific fluctuations in attentional experiences without the need to interrupt ongoing tasks. This would be useful given the demonstration by Seli, Carriere, et al. (2013) that probing too frequently during a task can lead to fewer reports of mind-wandering than when

probes are deployed less frequently. Of course, one solution is to reduce the frequency of immediate probes to reduce task disruption. Indeed, Welhaf et al. (2022) proposed that as few as eight thought probes might be sufficient for capturing basic mind-wandering reports during a task. However, reducing the frequency of probes has the knock-on effect of leaving an insufficient number of datapoints to faithfully capture the temporal fluctuations in attentional states. Thus, using retrospective probes instead of immediate probes, or supplementing sparse immediate-probe data with retrospective-probe data, might be a useful way forward in some experimental contexts.

The problem of immediate probes interfering with task engagement could also be solved by using aggregate retrospective probes, as has been done in the past. Along these lines, Smallwood and Schooler (2015) proposed that:

One solution [to having to probe participants too frequently in the moment] is to acquire self-report data after participants have completed an experimental session. Although this measure necessarily depends on memory, it allows the collection of data without artificial disruptions and is useful because by preserving the integrity of time-course data, it allows temporal properties in objective measures to be related to [experience sampling, i.e., immediate thought probing] data. (p. 495)

However, although aggregate retrospective measures such as the DSSQ might be sufficient for capturing individual differences in overall attentional propensities, they are not suitable for capturing the temporal dynamics of attention across the duration of a task. Thus, using stimulated recall to retrospectively probe individuals' RMAss provides a unique opportunity to index dynamic changes in attentional experiences after the fact, without the need to interrupt ongoing tasks.

## **Conclusion**

We found an impressive degree of overlap between participants' immediate and retrospective ratings when measuring attentional engagement, but much less so when measuring the more specific state of mind-wandering. Retrospective reports also correlated with memory for video content. These findings suggest that people may have enduring remembered momentary attentional states (RMASs) for general attentional engagement but less-robust memories for the specific experience of mind-wandering. The present work constitutes a first systematic exploration of RMASs and opens the door to this exciting new research area of human attention. Furthermore, the present work introduces a methodology that can be used in future explorations of RMASs and one that could be used instead of immediate thought probes in some cases.

## References

- Alaybek, B., Dalal, R. S., Fyffe, S., Aitken, J. A., Zhou, Y., Qu, X., Roman, A., & Baines, J. I. (2022). All's well that ends (and peaks) well? A meta-analysis of the peak-end rule and duration neglect. *Organizational Behavior and Human Decision Processes*, 170. <https://doi.org/10.1016/j.obhdp.2022.104149>
- Anderson, B. (2011). There is no such thing as attention. *Frontiers in Psychology*, 2, 246. <https://doi.org/10.3389/fpsyg.2011.00246>
- Anderson, T., Petranker, R., Lin, H., & Farb, N. A. S. (2020). The metronome response task for measuring mind wandering: Replication attempt and extension of three studies by Seli et al. *Attention, Perception, & Psychophysics*, 83(1), 315-330. <https://doi.org/10.3758/s13414-020-02131-x>
- Antrobus, J. S. (1968). Information theory and stimulus-independent thought. *British Journal of Psychology*, 59(4), 423-430. <https://doi.org/10.1111/j.2044-8295.1968.tb01157.x>
- Baird, B., Smallwood, J., Lutz, A., & Schooler, J. W. (2014). The decoupled mind: Mind-wandering disrupts cortical phase-locking to perceptual events. *Journal of Cognitive Neuroscience*, 26(11), 2596-2607. [https://doi.org/10.1162/jocn\\_a\\_00656](https://doi.org/10.1162/jocn_a_00656)
- Banks, J. B., Welhaf, M. S., & Srour, A. (2015). The protective effects of brief mindfulness meditation training. *Consciousness and Cognition*, 33, 277-285. <https://doi.org/10.1016/j.concog.2015.01.016>
- Barker, P. M., & McDougall, T. J. (2020). Two interpolation methods using multiply-rotated piecewise cubic hermite interpolating polynomials. *Journal of Atmospheric and Oceanic Technology*, 37(4), 605-619. <https://doi.org/10.1175/jtech-d-19-0211.1>
- Barron, E., Riby, L. M., Greer, J., & Smallwood, J. (2011). Absorbed in thought: The effect of mind wandering on the processing of relevant and irrelevant events. *Psychological Science*, 22(5), 596-601. <https://doi.org/10.1177/0956797611404083>
- Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., Gijsselaers, W., & Westendorp, J. (2006). Cognitive load measurements and stimulated recall interviews for studying the effects of information and communications technology. *Educational Technology Research and Development*, 56(3), 309-328. <https://doi.org/10.1007/s11423-006-9020-7>
- Bloom, B. S. (1953). Thought-processes in lectures and discussions. *The Journal of General Education*, 7(3), 160-169.
- Brosowsky, N.P., DeGutis, J., Esterman, M., Smilek, D., & Seli, P. (2020). Mind wandering, motivation, and task performance over time: Evidence that motivation insulates people from the negative effects of mind wandering. *Psychology of Consciousness: Theory, Research, and Practice* [online ahead of print]. <https://doi.org/10.1037/csn0000263>

- Brunsdon, T. M., & Skinner, C. J. (1987). The analysis of dependencies between time series in psychological experiments. *British Journal of Mathematical and Statistical Psychology*, 40(2), 125-139. <https://doi.org/https://doi.org/10.1111/j.2044-8317.1987.tb00873.x>
- Calderhead, J. (1981). Stimulated recall: A method for research on teaching. *British Journal of Educational Psychology*, 51(2), 211-217.
- Carriere, J. S., 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>
- Carriere, J. S. A., Seli, P., & Smilek, D. (2013). Wandering in both mind and body: Individual differences in mind wandering and inattention predict fidgeting. *Canadian Journal of Experimental Psychology*, 67(1), 19-31. <https://doi.org/10.1037/a0031438>
- 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*, 106(21), 8719-8724. <https://doi.org/10.1073/pnas.0900234106>
- Christoff, K., Irving, Z. C., Fox, K. C. R., Spreng, R. N., & Andrews-Hanna, J. R. (2016). Mind-wandering as spontaneous thought: A dynamic framework. *Nature Reviews Neuroscience*, 17(11), 718-731. <https://doi.org/10.1038/nrn.2016.113>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Csikszentmihalyi, M., & Csikszentmihalyi, I. S. (1992). *Optimal experience: Psychological studies of flow in consciousness*. Cambridge University Press.
- Dan, E. L., Dînşoreanu, M., & Mureşan, R. C. (2020). Accuracy of six interpolation methods applied on pupil diameter data. In *2020 IEEE International Conference on Automation, Quality and Testing, Robotics (AQTR)* (pp. 1-5). IEEE.
- Danckert, J. (2018). Special topic introduction: Understanding engagement: Mind-wandering, boredom and attention. *Experimental Brain Research*, 236(9), 2447-2449. <https://doi.org/10.1007/s00221-017-4914-7>
- Danckert, J., Hammerschmidt, T., Marty-Dugas, J., & Smilek, D. (2018). Boredom: Under-aroused and restless. *Consciousness and Cognition*, 61, 24-37. <https://doi.org/10.1016/j.concog.2018.03.014>
- Delorme, A., & Brandmeyer, T. (2019). When the meditating mind wanders. *Current Opinion in Psychology*, 28, 133-137. <https://doi.org/https://doi.org/10.1016/j.copsyc.2018.12.006>
- Demetriou, A., & Bakracevic, K. (2009). Reasoning and self-awareness from adolescence to middle age: Organization and development as a function of education. *Learning and Individual Differences*, 19(2), 181-194. <https://doi.org/10.1016/j.lindif.2008.10.007>

- Dempsey, N. P. (2010). Stimulated recall interviews in ethnography. *Qualitative Sociology*, 33(3), 349-367. <https://doi.org/10.1007/s11133-010-9157-x>
- Derrick, T. R., & Thomas, J. M. (2004). Time series analysis: The cross-correlation function. In N. Stergiou (Ed.), *Innovative Analysis of Human Movement: Analytical Tools for Human Movement Research* (pp. 189-205). Human Kinetics Publishers.
- Derryberry, D., & Reed, M. A. (2002). Anxiety-related attentional biases and their regulation by attentional control. *Journal of Abnormal Psychology*, 111(2), 225-236. <https://doi.org/10.1037//0021-843x.111.2.225>
- Dixon, P. & Li, H. (2013). Mind wandering in text comprehension under dual-task conditions. *Frontiers in Psychology*, 4, 1-14. <https://doi.org/10.3389/fpsyg.2013.00682>
- Ellison, W. D., Trahan, A. C., Pinzon, J. C., Gillespie, M. E., Simmons, L. M., & King, K. Y. (2020). For whom, and for what, is experience sampling more accurate than retrospective report? *Personality and Individual Differences*, 163, 110071. <https://doi.org/https://doi.org/10.1016/j.paid.2020.110071>
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149-1160. <https://doi.org/10.3758/BRM.41.4.1149>
- 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>
- Fritsch, F. N., & Carlson, R. E. (1980). Monotone piecewise cubic interpolation. *SIAM Journal on Numerical Analysis*, 17, 238-246. <https://doi.org/10.1137/0717021>
- Gass, S. M., & Mackey, A. (2016). *Stimulated recall methodology in applied linguistics and L2 research*. Routledge.
- Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition*, 4, 1-21.
- Gregersen, T., MacIntyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, 98(2), 574-588. <https://doi.org/10.1111/j.1540-4781.2014.12084.x>
- Gregson, R. A. M. (2014). *Time series in psychology*. Psychology Press.
- Hasenkamp, W., Wilson-Mendenhall, C. D., Duncan, E., & Barsalou, L. W. (2012). Mind wandering and attention during focused meditation: A fine-grained temporal analysis of fluctuating cognitive states. *Neuroimage*, 59(1), 750-760. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2011.07.008>

- Heath, R. A. (2000). *Nonlinear dynamics: Techniques and applications in psychology* (1st ed.). Psychology Press. [https://doi.org/https://doi.org/10.4324/9781410605702](https://doi.org/10.4324/9781410605702)
- Kahaner, D., Moler, C., & Nash, S. (1988). *Numerical methods and software*. Prentice Hall.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when. *Psychological Science, 18*(7), 614-621.
- Kane, M. J., Meier, M. E., Smeekens, B. A., Gross, G. M., Chun, C. A., Silvia, P. J., & Kwapil, T. R. (2016). Individual differences in the executive control of attention, memory, and thought, and their associations with schizotypy. *Journal of Experimental Psychology: General, 145*(8), 1017-1048. <https://doi.org/10.1037/xge0000184>
- Kane, M. J., Smeekens, B. A., Meier, M. E., Welhaf, M. S., & Phillips, N. E. (2021). Testing the construct validity of competing measurement approaches to probed mind-wandering reports. *Behavioral Research Methods, 53*(6), 2372-2411. <https://doi.org/10.3758/s13428-021-01557-x>
- Killingsworth, M. A., & Gilbert, D. T. (2010). A wandering mind is an unhappy mind. *Science, 330*, 932. <https://doi.org/10.1126/science.1192439>
- Kleiner, D. E. E. (2009). *Lifestyles of the rich and famous: Houses and villas at Pompeii*, Open Yale Courses. <https://oyc.yale.edu/history-art/hsar-252/lecture-5>
- Kline, R. B. (1998). *Principles and practice of structural equation modelling*. Guilford Press.
- Kopp, K., Mills, C., & D'Mello, S. (2016). Mind wandering during film comprehension: The role of prior knowledge and situational interest. *Psychonomic Bulletin & Review, 23*(3), 842-848. <https://doi.org/10.3758/s13423-015-0936-y>
- Lee, I. A., & Preacher, K. J. (2013). *Calculation for the test of the difference between two dependent correlations with one variable in common*. In <http://quantpsy.org>
- Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal, 29*(6), 861-878. <https://doi.org/10.1080/0141192032000137349>
- Marty-Dugas, J., Howes, L., & Smilek, D. (2021). Sustained attention and the experience of flow. *Psychological Research, 85*(7), 2682-2696. <https://doi.org/10.1007/s00426-020-01433-x>
- Matthews, G., Joyner, L., Gilliland, K., S.Campbell, Falconer, S., & Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. *Personality Psychology in Europe, 7*, 335-350.
- Matthews, G., Szalma, J., Panganiban, A. R., Neubauer, C., & Warm, J. S. (2013). Profiling task stress with the Dundee Stress State Questionnaire. *Psychology of Stress: New Research, 1*, 49-90.

- 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 & Cognition*, 35(1), 196-204. <https://doi.org/10.1037/a0014104>
- McVay, J. C., & Kane, M. J. (2012). Drifting from slow to "D'oh!": Working memory capacity and mind wandering predict extreme reaction times and executive control errors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38(3), 525-549. <https://doi.org/10.1037/a0025896>
- 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>
- Meade, P., & McMeniman, M. (1992). Stimulated recall: An effective methodology for examining successful teaching in science. *The Australian Educational Researcher*, 19(3), 1-18.
- Morgan, A. (2007). Using video-stimulated recall to understand young children's perceptions of learning in classroom settings. *European Early Childhood Education Research Journal*, 15(2), 213-226. <https://doi.org/10.1080/13502930701320933>
- Myin-Germeys, I., Oorschot, M., Collip, D., Lataster, J., Delespaul, P., & van Os, J. (2009). Experience sampling research in psychopathology: Opening the black box of daily life. *Psychological Medicine*, 39(9), 1533-1547. <https://doi.org/10.1017/S0033291708004947>
- O'Brien, J. (1993). Action research through stimulated recall. *Research in Science Education*, 23(1), 214-221.
- Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, 46(4), 1023-1031.
- Phillips, N. E., Mills, C., D'Mello, S., & Risko, E. F. (2016). On the influence of re-reading on mind wandering. *Quarterly Journal of Experimental Psychology*, 69(12), 2338-2357. <https://doi.org/10.1080/17470218.2015.1107109>
- Ralph, B. C., Onderwater, K., Thomson, D. R., & Smilek, D. (2017). Disrupting monotony while increasing demand: Benefits of rest and intervening tasks on vigilance. *Psychological Research*, 81(2), 432-444. <https://doi.org/10.1007/s00426-016-0752-7>
- Randall, J. G., Oswald, F. L., & Beier, M. E. (2014). Mind-wandering, cognition, and performance: A theory-driven meta-analysis of attention regulation. *Psychological Bulletin*, 140(6), 1411-1431. <https://doi.org/10.1037/a0037428>
- Redelmeier, D. A., & Kahneman, D. (1996). Patients' memories of painful medical treatments: Real-time and retrospective evaluations of two minimally invasive procedures. *Pain*, 66(1), 3-8. [https://doi.org/https://doi.org/10.1016/0304-3959\(96\)02994-6](https://doi.org/https://doi.org/10.1016/0304-3959(96)02994-6)

- Ricker, T. J., Vergauwe, E., & Cowan, N. (2016). Decay theory of immediate memory: From Brown (1958) to today (2014). *Quarterly Journal of Experimental Psychology*, *69*(10), 1969-1995. <https://doi.org/10.1080/17470218.2014.914546>
- Risko, E. F., Anderson, N., Sarwal, A., Engelhardt, M., & Kingstone, A. (2012). Everyday attention: Variation in mind wandering and memory in a lecture. *Applied Cognitive Psychology*, *26*(2), 234-242. <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/https://doi.org/10.1016/j.compedu.2013.05.001>
- Rober, M. (2018). *The Super Mario Effect: Tricking your brain into learning more*, TEDxPenn. [https://www.ted.com/talks/mark\\_rober\\_the\\_super\\_mario\\_effect\\_tricking\\_your\\_brain\\_into\\_learning\\_more?language=en](https://www.ted.com/talks/mark_rober_the_super_mario_effect_tricking_your_brain_into_learning_more?language=en)
- Robison, M. K., Miller, A. L., & Unsworth, N. (2019). Examining the effects of probe frequency, response options, and framing within the thought-probe method. *Behavioral Research Methods*, *51*(1), 398-408. <https://doi.org/10.3758/s13428-019-01212-6>
- Seli, P., Beaty, R. E., Cheyne, J. A., Smilek, D., Oakman, J., & Schacter, D. L. (2018). How pervasive is mind wandering, really? *Consciousness and Cognition*, *66*, 74-78. <https://doi.org/10.1016/j.concog.2018.10.002>
- 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*, 430-430. <https://doi.org/10.3389/fpsyg.2013.00430>
- Seli, P., Cheyne, J. A., & Smilek, D. (2013). Wandering minds and wavering rhythms: Linking mind wandering and behavioral variability. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 1-5. <https://doi.org/10.1037/a0030954>
- 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 Science*, *20*(8), 605-617. <https://doi.org/10.1016/j.tics.2016.05.010>
- Shaw, G. A., & Giambra, L. (1993). Task-unrelated thoughts of college students diagnosed as hyperactive in childhood. *Developmental Neuropsychology*, *9*(1), 17-30. <https://doi.org/10.1080/87565649309540541>
- Sinnott, C., Kelly, M. A., & Bradley, C. P. (2017). A scoping review of the potential for chart stimulated recall as a clinical research method. *BMC Health Services Research*, *17*(1), 583. <https://doi.org/10.1186/s12913-017-2539-y>

- Smallwood, J., Davies, J. B., Heim, D., Finnigan, F., Sudberry, M., O'Connor, R., & Obonsawin, M. (2004). Subjective experience and the attentional lapse: Task engagement and disengagement during sustained attention. *Consciousness and Cognition, 13*(4), 657-690. <https://doi.org/10.1016/j.concog.2004.06.003>
- 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., Riby, L., Heim, D., & Davies, J. B. (2006). Encoding during the attentional lapse: Accuracy of encoding during the semantic sustained attention to response task. *Consciousness and Cognition, 15*(1), 218-231. <https://doi.org/10.1016/j.concog.2005.03.003>
- Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132*(6), 946-958. <https://doi.org/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*, 487-518. <https://doi.org/10.1146/annurev-psych-010814-015331>
- Soemer, A., & Schiefele, U. (2019). Text difficulty, topic interest, and mind wandering during reading. *Learning and Instruction, 61*, 12-22. <https://doi.org/10.1016/j.learninstruc.2018.12.006>
- Stawarczyk, D. (2018). Phenomenological properties of mind-wandering and daydreaming: A historical overview and functional correlates. In K. C. R. Fox & K. Christoff (Eds.), *The Oxford Handbook of Spontaneous Thought: Mind-Wandering, Creativity, and Dreaming* (pp. 193-214). Oxford University Press.
- 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>
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*(2), 245-251.
- 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, 110*(16), 6313-6317. <https://doi.org/10.1073/pnas.1221764110>
- Szpunar, K. K., Moulton, S., & Schacter, D. L. (2013). Mind wandering and education: From the classroom to online learning [Review]. *Frontiers in Psychology, 4*, 495. <https://doi.org/10.3389/fpsyg.2013.00495>

- Thomson, D. R., Besner, D., & Smilek, D. (2015). A resource-control account of sustained attention: Evidence from mind-wandering and vigilance paradigms. *Perspectives on Psychological Science*, *10*(1), 82-96. <https://doi.org/10.1177/1745691614556681>
- 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>
- Tusche, A., Smallwood, J., Bernhardt, B. C., & Singer, T. (2014). Classifying the wandering mind: Revealing the affective content of thoughts during task-free rest periods. *Neuroimage*, *97*, 107-116. <https://doi.org/https://doi.org/10.1016/j.neuroimage.2014.03.076>
- Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective, and Behavioral Neuroscience*, *16*(4), 601-615. <https://doi.org/10.3758/s13415-016-0417-4>
- Varao-Sousa, T.L. & Kingstone, A. (2019). 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*, *51*, 235-242. <https://doi.org/10.3758/s13428-018-1073-0>
- 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., Ralph, B. C. W., Mills, C., Bosch, N., Duncan, T. L., & Smilek, D. (2019). Disengagement during lectures: Media multitasking and mind wandering in university classrooms. *Computers & Education*, *132*, 76-89. <https://doi.org/https://doi.org/10.1016/j.compedu.2018.12.007>
- 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/10.1037/h0101808>
- Ward, A. F., & Wegner, D. M. (2013). Mind-blanking: When the mind goes away. *Frontiers in Psychology*, *4*, 650. <https://doi.org/10.3389/fpsyg.2013.00650>
- Weinstein, Y. (2018). Mind-wandering, how do I measure thee with probes? Let me count the ways. *Behavioral Research Methods*, *50*(2), 642-661. <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>

Welhaf, M. S., Meier, M. E., Smeekens, B. A., Silvia, P. J., Kwapil, T. R., & Kane, M. J. (2022). A "goldilocks zone" for mind wandering reports? A secondary data analysis of how few thought probes are enough for reliable and valid measurement. *Behavioral Research Methods*. <https://doi.org/https://doi.org/10.3758/s13428-021-01766-4>

Zeinali, M., Shahmorad, S., & Mirnia, K. (2014). Hermite and piecewise cubic Hermite interpolation of fuzzy data. *Journal of Intelligent & Fuzzy Systems*, 26, 2889-2898. <https://doi.org/10.3233/IFS-130955>

**Appendices**  
**Appendix A**  
**Post-Task Quizzes**

*\*Correct responses are italicized*

**More-Engaging Video (Super Mario Effect)**

1. Mark had his YouTube followers complete a puzzle that was designed to:
  - a. Test video game playing ability
  - b. Prove that anyone can learn how to code
  - c. *Compare perseverance on a task after losing points or not losing points (0:58)*
  - d. Award the fastest programmer with \$200
2. In the computer programming puzzle, participants who were not penalized for failed attempts on the puzzle were \_\_\_\_\_ than participants who were penalized 5 points
  - a. *More successful (2:12)*
  - b. Less successful
  - c. Equally successful
3. What is the Super Mario Effect?
  - a. Video games improve learning
  - b. Focusing on failures improves learning
  - c. *Sticking with a task improves learning (4:30)*
  - d. Video games improve hand-eye coordination
4. What does Mark do for a living?
  - a. Video game programmer
  - b. School teacher
  - c. Author
  - d. *Science YouTuber (5:00)*
5. Out of the projects that Mark has spoken about, which one took him the longest to create?
  - a. World's largest Nerf gun
  - b. World's largest super soaker
  - c. *Moving dart board (5:55)*
  - d. Snowball machine gun

### Less-Engaging Video (Roman History)

1. Most Roman residential houses were \_\_\_\_ high
  - a. *1 story (1:35)*
  - b. 2 stories
  - c. 3 stories
  - d. 4 stories
  
2. The large second-story windows were likely influenced by \_\_\_\_\_ architecture
  - a. Italian
  - b. French
  - c. *Greek (2:42)*
  - d. American
  
3. What material were shrines usually made of?
  - a. Stone
  - b. *Wood (10:12)*
  - c. Metal
  - d. Glass
  
4. Which of the following residential rooms was NOT discussed?
  - a. Bedroom
  - b. Dining room
  - c. Entrance
  - d. *Bathroom*
  
5. Roman houses and \_\_\_\_\_ were built very similarly
  - a. *Tombs (14:00)*
  - b. Outhouses
  - c. Cafés
  - d. Churches

## Appendix B

### Experiment 2: Mind-Wandering Cross-Correlation Analyses (Before Listwise Deletions)

As many of the participants' cross-correlation data in Experiment 1 were deleted listwise due to having zero variability within either their immediate or retrospective probe reports for at least one of the video conditions (39/102 removed; Figure 5 in main text), we sought to examine whether the cross-correlation results were affected by these removals. Here, we included all possible data, only excluding the specific cross-correlation coefficients that were impossible to calculate (e.g., keeping a participant's less-engaging video's cross-correlation if it was computable even if they had no corresponding more-engaging video cross-correlation). This resulted in 87 participants' data being included for the less-engaging video condition and 68 participants' data being included for the more-engaging video condition. As paired less-engaging and more-engaging data were required to compare the two video types, we only re-ran the null-effect and peakiness analyses on the available data.

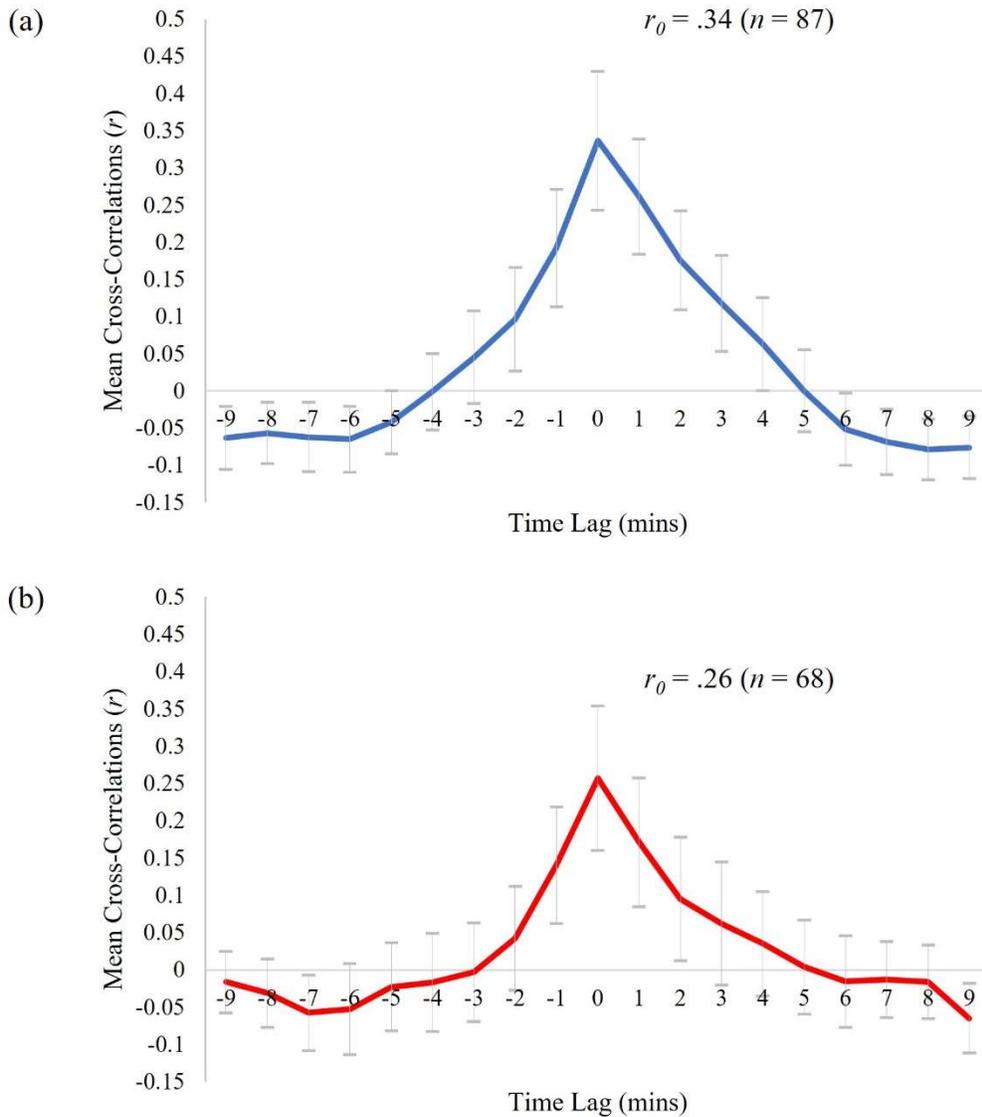
**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients of attentional engagement significantly differed from 0 for both the less-engaging ( $M = .34$ ,  $SD = .44$ ; one-sample  $t[86] = 7.14$ ,  $p < .001$ ,  $d_z = .77$ ; Figure B1a) and more-engaging video conditions ( $M = .26$ ,  $SD = .40$ ; one-sample  $t[67] = 5.28$ ,  $p < .001$ ,  $d_z = .64$ ; Figure B1b). Therefore, participants rated their attentional engagement similarly when probed immediately and retrospectively, demonstrating that specific prior moments of attentional engagement can be accurately probed after the fact even when participants were not deleted listwise.

The 95% confidence intervals in Figures B1a and B1b suggest that the cross-correlation coefficients were highest around the 0 lag for both video conditions, as they were significantly

higher than time lags -9 to -2 and +3 to +9. Therefore, the congruence between participants' immediate and retrospective reports of attentional engagement was strongest when the timeseries were directly overlaid and/or only shifted by up to 2 minutes, suggesting that participants were reporting on their RMASs of attentional engagement even when participants were not deleted listwise.

**Figure B1.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging and (b) More-Engaging Video Conditions of Experiment 1 (Before Listwise Deletions).



Note. Error bars represent 95% confidence intervals.

## Appendix C

### Experiment 2: Mind-Wandering Cross-Correlation Analyses (Before Listwise Deletions)

As many of the participants' cross-correlation data in Experiment 2 were deleted listwise due to having zero variability within either their immediate or retrospective probe reports for at least one of the mind-wandering types in at least one of the video conditions (39/97 removed; Figure 7 in main text), we sought to examine whether the cross-correlation results were affected by these removals. Here, we included all possible data, only excluding the specific cross-correlation coefficients that were impossible to calculate (e.g., keeping a participant's less-engaging video's cross-correlation if it was computable even if they had no corresponding more-engaging video cross-correlation). This resulted in 79 participants' intentional data being included for the less-engaging video condition, 86 participants' unintentional data being included for the less-engaging video condition, 67 participants' intentional data being included for the more-engaging condition, and 74 participants' unintentional data being included for the more-engaging video condition. As paired less-engaging and more-engaging data were required to compare the two video types, we only re-ran the null-effect and peakiness analyses on the available data.

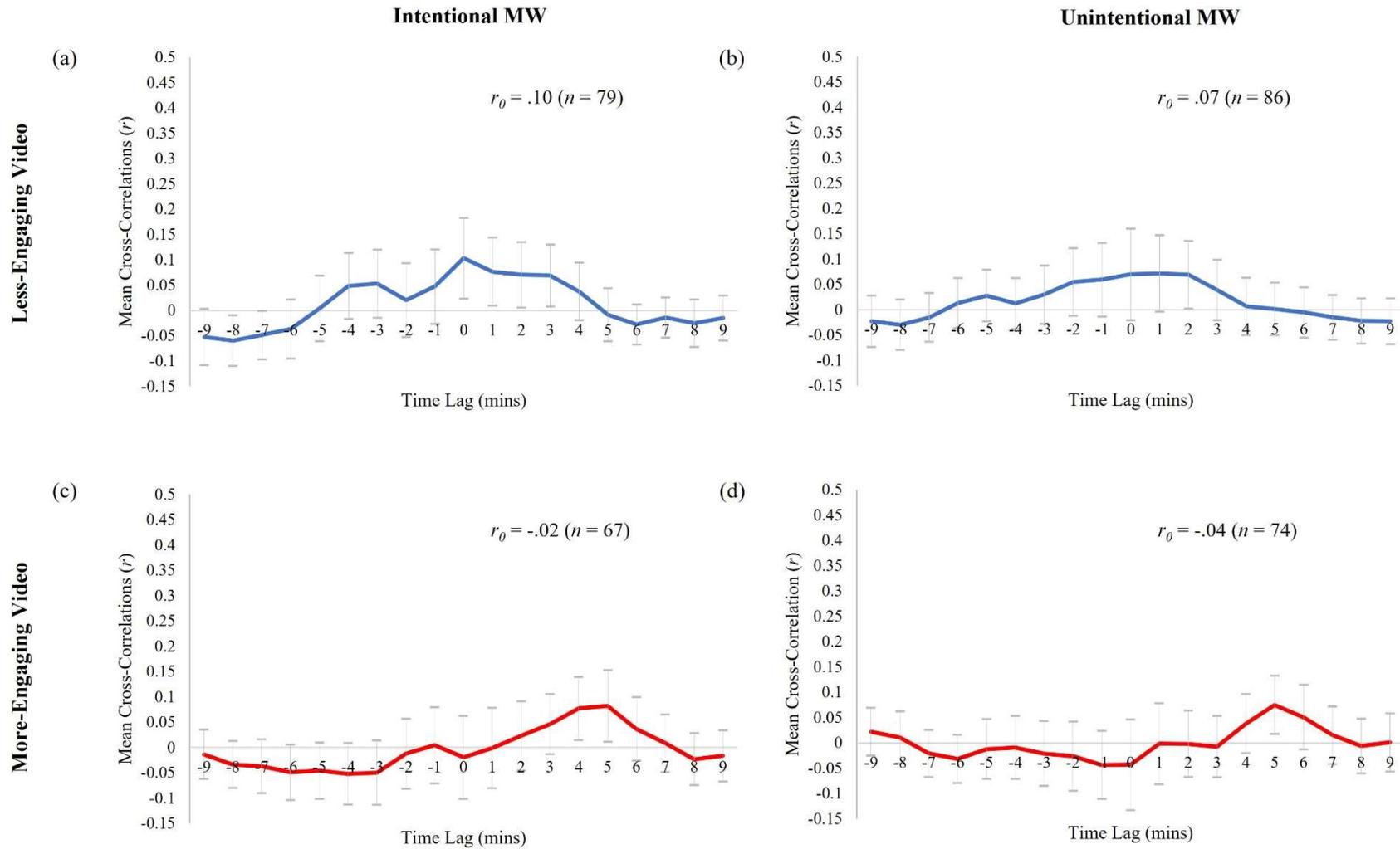
**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients only differed significantly from 0 for the less-engaging video's intentional probes ( $M = .10$ ,  $SD = .36$ ; one-sample  $t[78] = 2.56$ ,  $p = .013$ ,  $d_z = .29$ ; Figure C1a) but not for the less-engaging video's unintentional probes ( $M = .07$ ,  $SD = .42$ ; one-sample  $t[85] = 1.54$ ,  $p = .129$ ,  $d_z = .17$ ; Figure C1b). Furthermore, the average cross-correlation coefficients for the more-engaging video condition did not differ significantly from 0 for either the intentional probes ( $M = -.02$ ,  $SD = .34$ ; one-sample  $t[66] = -.47$ ,  $p = .637$ ,  $d_z = -.06$ ; Figure C1c) nor the

unintentional probes ( $M = -.04$ ,  $SD = .39$ ; one-sample  $t[73] = -.97$ ,  $p = .335$ ,  $d_z = -.11$ ; Figure C1d). Therefore, participants only rated their mind-wandering similarly when probed immediately and retrospectively about their intentional mind-wandering during one video condition, demonstrating that mind-wandering might not be accurately probed after the fact even when participants were not deleted listwise.

The 95% confidence intervals in Figures C1a-d suggest that the cross-correlation coefficients at the 0 lag were not significantly higher than at any other time lag, except for the less-engaging intentional mind-wandering condition (Figure C1a) where the 0 lag was significantly higher than lags -7 to -9 and +6 to +9. Therefore, the overall congruence between participants' immediate and retrospective reports of mind-wandering were not strongest when directly overlaid, suggesting that participants might not have been reporting on their RMASs of mind-wandering or that they had poor recollection of those states even when participants were not deleted listwise.

**Figure C1.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging Intentional MW, (b) Less-Engaging Unintentional MW, (c) More-Engaging Intentional MW, & (d) More-Engaging Unintentional MW Conditions of Experiment 2 (Before Listwise Deletions).



Note. Error bars represent 95% confidence intervals.

## Appendix D

### Experiment 2 Cross-Correlation Results

#### (Combined Overall Mind-Wandering)

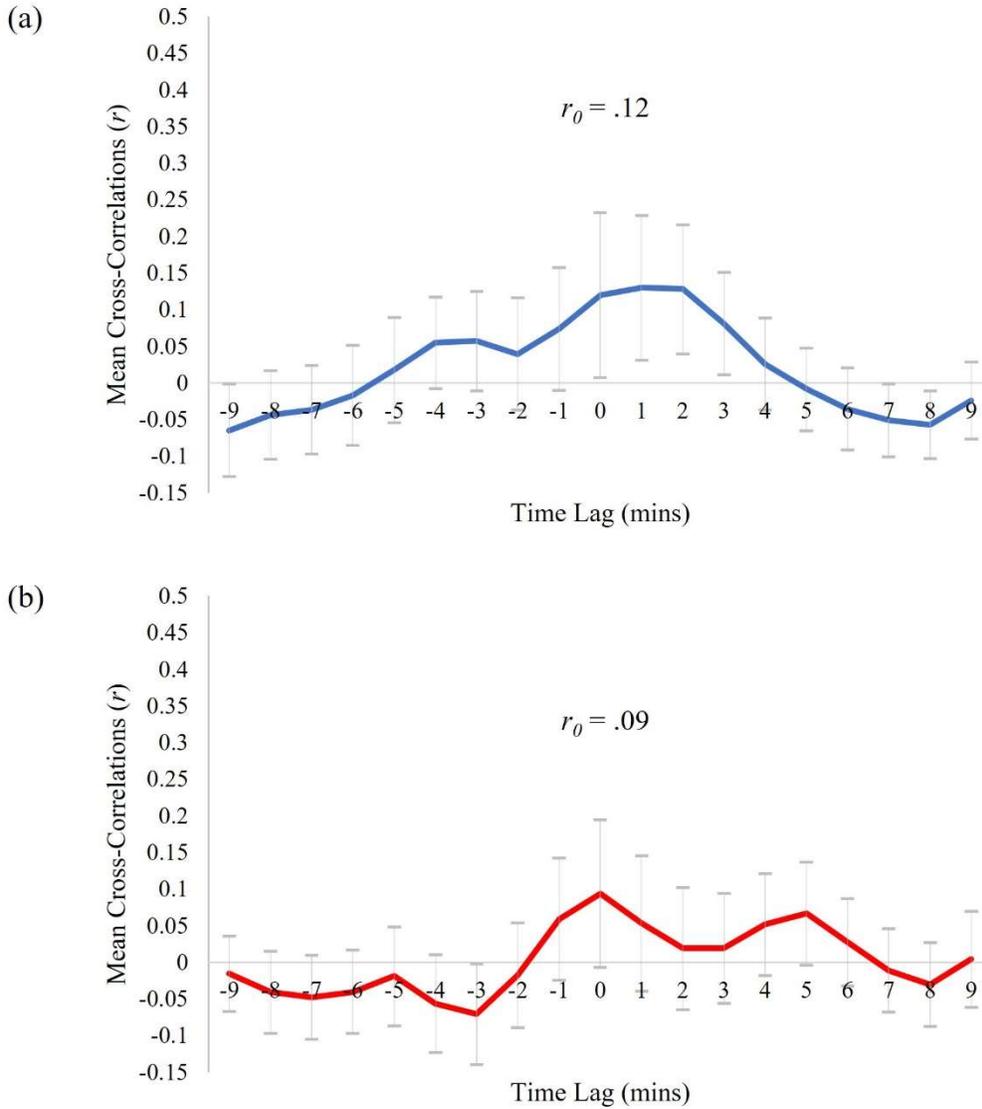
**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients of overall combined mind-wandering only differed significantly from 0 for the less-engaging video condition ( $M = .12$ ,  $SD = .43$ ; one-sample  $t[57] = 2.14$ ,  $p = .037$ ,  $d_z = .28$ ; Figure D1a) but not for the more-engaging video condition ( $M = .09$ ,  $SD = .38$ ; one-sample  $t[57] = 1.86$ ,  $p = .068$ ,  $d_z = .24$ ; Figure D1b). This effect was driven by the intentional mind-wandering reports. Therefore, participants only rated their overall mind-wandering similarly when probed immediately and retrospectively in one video condition but not in another, demonstrating that mind-wandering might only be accurately probed after the fact in certain contexts.

As in the intentional and unintentional analyses, the 95% confidence intervals in Figures D1a and D1b suggest that the cross-correlation coefficients at the 0 lag for combined mind-wandering were not significantly higher than at any other time lag, except for the less-engaging condition (Figure D1a) where the 0 lag was significantly higher than lags -9, +7 and +8. This effect was driven by the intentional mind-wandering reports. Therefore, the congruence between participants' immediate and retrospective reports of overall mind-wandering were not strongest when directly overlaid, again suggesting that participants might not have been reporting on their RMASs of mind-wandering or that they had poor recollection of those states.

**Comparing video conditions.** Finally, there was not a significant main effect of video type (less-engaging versus more-engaging) on mean cross-correlation coefficients at a time lag of 0 ( $t[57] = .360$ ,  $p = .721$ ). Therefore, the video condition did not affect participants' concordance between the immediate and retrospective reports of overall mind-wandering.

**Figure D1.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging Combined MW and (b) More-Engaging Combined MW Conditions of Experiment 2 (N = 58).



Note. MW = mind-wandering. Error bars represent 95% confidence intervals.

The same results were observed when all possible participants were included in the analyses (i.e., before listwise deletions).

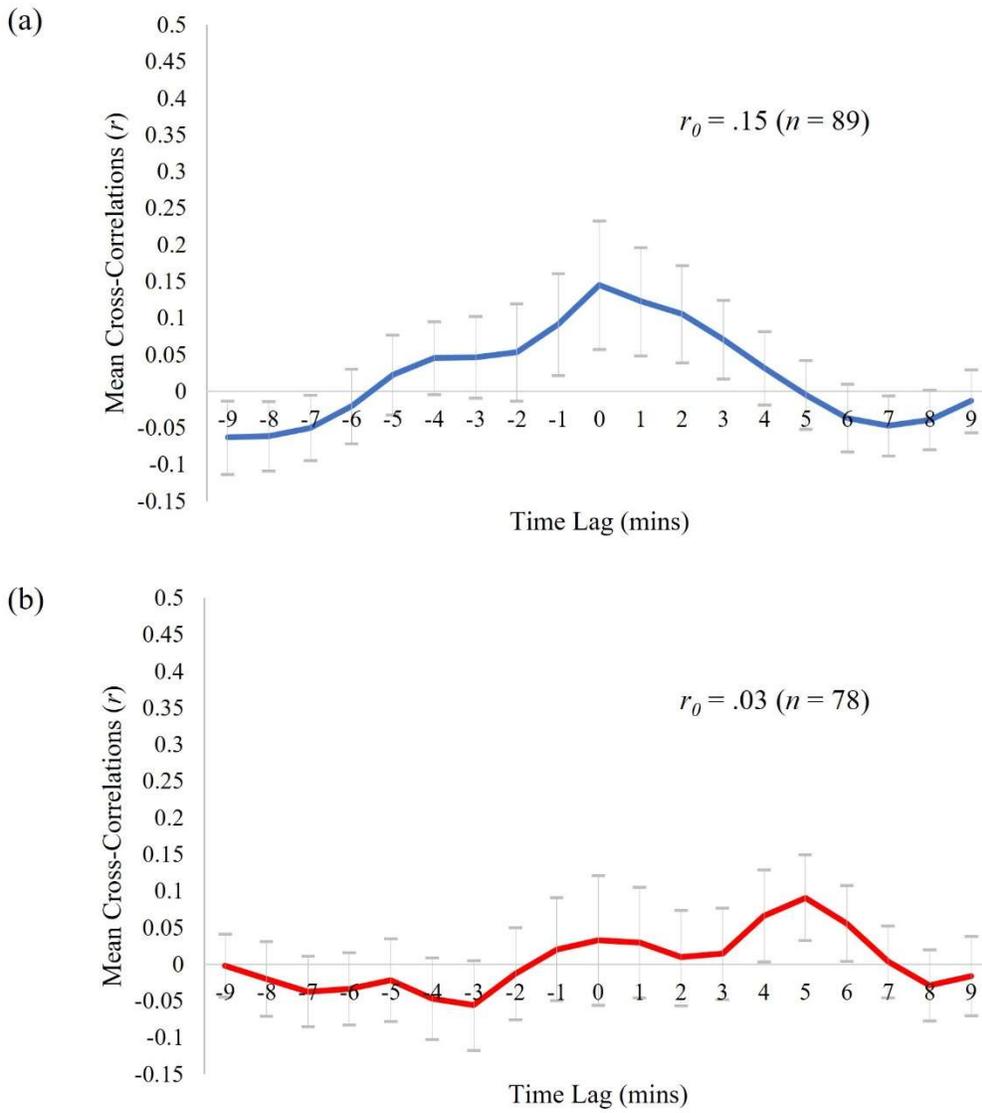
**Evaluating concordance between immediate and retrospective reports.** Average cross-correlation coefficients of overall combined mind-wandering only differed significantly from 0 for the less-engaging video condition ( $M = .15$ ,  $SD = .42$ ; one-sample  $t[88] = 3.29$ ,  $p =$

.001,  $d_z = .35$ ; Figure D2a) but not for the more-engaging video condition ( $M = .03$ ,  $SD = .39$ ; one-sample  $t[77] = .75$ ,  $p = .458$ ,  $d_z = .08$ ; Figure D2b). This effect was driven by the intentional mind-wandering reports. Therefore, participants only rated their overall mind-wandering similarly when probed immediately and retrospectively in one video condition but not in another, demonstrating that mind-wandering might only be accurately probed after the fact in certain contexts.

As in the intentional and unintentional analyses, the 95% confidence intervals in Figures D2a and D2b suggest that the cross-correlation coefficients at the 0 lag for combined mind-wandering were not significantly higher than at any other time lag, except for the less-engaging condition (Figure D2a) where the 0 lag was significantly higher than lags -6 to -9 and +5 to +9. This effect was driven by the intentional mind-wandering reports. Therefore, the congruence between participants' immediate and retrospective reports of overall mind-wandering were not strongest when directly overlaid, again suggesting that participants might not have been reporting on their RMASs of mind-wandering or that they had poor recollection of those states even when participants were not deleted listwise.

**Figure D2.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Less-Engaging Combined MW and (b) More-Engaging Combined MW Conditions (Before Listwise Deletions).



Note. Error bars represent 95% confidence intervals.

## Appendix E

### Comparing Attentional Engagement and Mind-Wandering Cross-Correlation Analyses (Before Listwise Deletions)

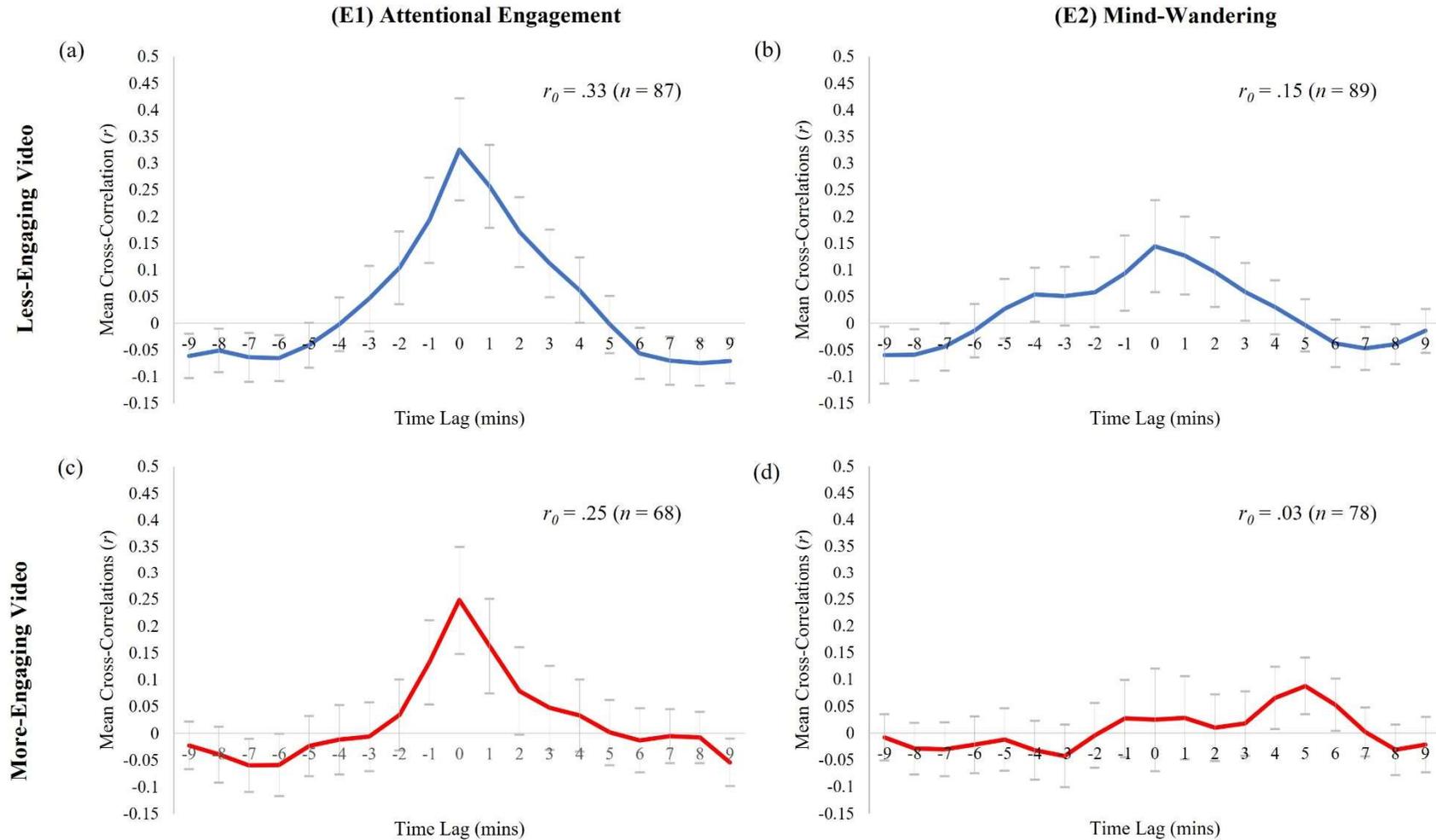
As many of the participants' cross-correlation data in Experiment 1 and Experiment 2 were deleted listwise due to having zero variability within either their immediate or retrospective probe reports for at least one of the video conditions (Figure 8 in main text), we sought to examine whether the cross-correlation results were affected by these removals. Here, we included all possible data, only excluding the specific cross-correlation coefficients that were impossible to calculate (e.g., keeping a participant's less-engaging video's cross-correlation if it was computable even if they had no corresponding more-engaging video cross-correlation). This resulted in 87 participants' data being included for the less-engaging video condition of Experiment 1, 89 participants' data being included for the less-engaging video condition of Experiment 2, 68 participants' data being included for the more-engaging video condition of Experiment 1, and 78 participants' data being included for the more-engaging video condition of Experiment 2.

To compare the cross-correlation results from Experiment 1 (attentional engagement) and Experiment 2 (overall mind-wandering), we standardized each participant's raw probe responses, re-ran the cross-correlation analyses, then compared the results using independent *t*-tests. For the less-engaging video condition, average cross-correlation coefficients were significantly higher at a time lag of 0 when asking participants about attentional engagement ( $M = .33$ ,  $SD = .45$ ; Figure E1a) than when asking about overall mind-wandering ( $M = .15$ ,  $SD = .41$ ; Figure E1b),  $t(174) = 2.79$ ,  $p = .006$ ,  $d = .42$ . The same was observed for the more-engaging video condition: average cross-correlation coefficients were significantly higher at a time lag of 0 for attentional engagement ( $M = .25$ ,  $SD = .42$ ; Figure E1c) than for overall mind-wandering ( $M = .03$ ,  $SD =$

.42; Figure E1d),  $t(144) = 3.20$ ,  $p = .002$ ,  $d = .53$ . Therefore, there was stronger concordance between the immediate and retrospective reports of attention when participants were probed about levels of attentional engagement versus mind-wandering, even when participants were not deleted listwise.

**Figure E1.**

Mean Cross-Correlation Scores as a Function of Time Lag from the (a) Standardized E1 Less-Engaging, (b) Standardized E2 Less-Engaging, (c) Standardized E1 More-Engaging and (d) Standardized E2 More-Engaging Conditions (Before Listwise Deletions).



Note. E1 = Experiment 1 (attentional engagement probes), E2 = Experiment 2 (overall mind-wandering probes). Error bars represent 95% confidence intervals.